

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 00:46:46 ; Search time 3592 Seconds
(without alignments)
12863.509 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtctctgtggcctcct.....ggtctctgagggcgtgtgta 2853

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 24

Total number of hits satisfying chosen parameters: 15

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpi: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_fod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
C 1	53	1.9	544	12	BF078689	229090 MA
C 2	39	1.4	195	10	AW935807	QV3-DT001
C 3	38	1.3	528	17	AQ809642	HS-4758.A
C 4	36	1.3	961	10	BB612189	BB612189
C 5	35	1.2	469	17	AZ842795	2M0141119
C 6	35	1.2	522	10	BE553572	ur47b04.y

```

7      30      1.1      305      12      BF320986      BF320986      uz57h10.Y
8      29      1.0      400      17      AQ440250      AQ440250      HS_5071.B
9      28      1.0      415      10      AW485848      AW485848      68315 MAR
10     27      0.9      169      17      B01986       B01986       CSRL-145A3-
11     26      0.9      627      10      BE281880      BE281880      601099657
12     25      0.9      247      10      BB604671      BB604671      BB604671
13     25      0.9      849      13      BI103177      BI103177      602889417
14     24      0.8      464      10      BE566088      BE566088      149119 MA
15     24      0.8      664      13      BI692986      BI692986      603344521

```

ALIGNMENTS

```

RESULT 1
BF078689/c
LOCUS      BF078689      544 bp      mRNA      linear      EST 18-OCT-2000
DEFINITION 229090 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION  BF078689
VERSION    BF078689.1 GI:10872519
KEYWORDS   EST.
SOURCE     pig.
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 544)
AUTHORS   Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
            Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Langreid, W.W.
            and Keele, J.W.
TITLE      Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
JOURNAL    Unpublished (2000)
COMMENT    Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 43 row: P column: 17
Seq primer: ATTAGGTGACACTATAG.
FEATURES   Location/Qualifiers
            1..544
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone_lib="MARC 2PIG"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /note="vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
            Library made from pooled tissue from testis, ovary,
            endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 117 a 154 c 169 g 104 t
ORIGIN

```

source

```

Query Match      1.9%; Score 53; DB 12; Length 544;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 ACGGGCGGCGACTCTACCGGCATCCAGCATCTCTCAACCCCATCACATCGT 794
      |||||
DB 361 ACGGGCGGCGACTCTACCGGCATCCAGCATCTCTCAACCCCATCACATCGT 309

```

```

RESULT 2
AW935807/c
LOCUS      AW935807      195 bp      mRNA      linear      EST 30-MAY-2000
DEFINITION QV3-DT0019-081299-039-c08 DT0019 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW935807

```

```

VERSION      AW935807.1  GI:81111213
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE    1 (bases 1 to 195)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
             Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
             Goldman,G.H., Calvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
             Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
             M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
TITLE        Shotgun sequencing of the human transcriptome with ORF expressed
             sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-27047001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the PAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=et2-QV3-DT0019-081
             299-039-c08&t3=1999-12-08&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 58
             High quality sequence stop: 109.
FEATURES     Location/Qualifiers
             1..195
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="DT0019"
             /dev_stage="Adult"
             /note="Organ: denis_drash; Vector: puc18; Site_1: SmaI;
             Site_2: SmaI; A mini-library was made by cloning products
             derived from ORESTES PCR (U.S. Letters Patent application
             No. 196,716 - Ludwig Institute for Cancer Research)
             profiles into the puc 18 vector. Reverse transcription of
             tissue mRNA and cDNA amplification were performed under
             low stringency conditions."
BASE COUNT   37 a 60 c 52 g 46 t
ORIGIN
Query Match 1.4%; Score 39; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1542 GTGGATGTTCTCTGGCCCAATGGATGCCCTATGGCCC 1580
      |||||||||||||||||||||||||||||||||||||||
Db 155 GGTGGATGTTCTCTGGCCCAATGGATGCCCTATGGCCC 117

RESULT 3
A0809642
LOCUS       HS_4758_AL_D11_77A CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=4758 Col=21 Row=G, DNA sequence.
ACCESSION  A0809642
VERSION     A0809642.1 GI:5728884
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE    1 (bases 1 to 528)
AUTHORS      Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
             Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
             Hood,L.

```

```

TITLE        Sequence-tagged connectors: A sequence approach to mapping and
             scanning the human genome
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE      99380589
COMMENT      Contact: Mahairas GG, Wallace JC, Hood L
             High Throughput Sequencing Center
             University of Washington
             401 Queen Anne Avenue North, Seattle, WA 98109, USA
             Tel: (206) 616-3618
             Fax: (206) 616-3887
             Email: jwallace@u.washington.edu
             Clones may be purchased from Research Genetics (info@resgen.com).
             BAC end Web Server: http://www.htsc.washington.edu
             Plate: 4758 row: G column: 21
             Seq primer: T7
             Class: BAC ends
             High quality sequence stop: 528.
FEATURES     Location/Qualifiers
             1..528
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="CIT Approved Human Genomic Sperm Library D"
             /sex="male"
             /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
             E-Coli DH10B"
BASE COUNT   138 a 145 c 100 g 142 t 3 others
ORIGIN
Query Match 1.3%; Score 38; DB 17; Length 528;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2114 GGCCTCAAGCATCGACATCGCCGCGGCTTACAAA 2151
      |||||||||||||||||||||||||||||||||||||||
Db 296 GGCCTCAAGCATCGACATCGCCGCGGCTTACAAA 333

RESULT 4
BB612189
LOCUS       BB612189 RIKEN full-length enriched, 0 day neonate skin Mus
DEFINITION musculus cDNA clone 4631401M01 5', mRNA sequence.
ACCESSION  BB612189
VERSION     BB612189.1 GI:16453236
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 961)
AUTHORS      Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
             Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
             M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Sasaki
             Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
             D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
             Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
             Muramatsu,M. and Hayashizaki,Y.
             RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
             Unpublished (2001)
TITLE        Yoshihide Hayashizaki
JOURNAL      Laboratory for Genome Exploration Research Group, RIKEN Genomic
COMMENT      Sciences Center (GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@sc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/
             Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
             M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.,
             Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new

```


REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 469)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hsu,H., C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: gdunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0141 row: I column: 19
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 469.

FEATURES

Location/Qualifiers

1. .469

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0141119"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN.

Query Match 1.2%; Score 35; DB 17; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy	1942	GTCTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTG	1978
Db	327	GTCTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTG	293

RESULT 6
BE553572

ur47b04.y1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:3153391 5' similar to TR:O54768 O54768 A DISINTEGRIN AND METALLOPROTEASE WITH THROMBOSPONDIN MOTIFS ;, mRNA sequence.

ACCESSION

VERSION
KEYWORDS

SOURCE
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 522)

REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1056147

Seq primer: -40RP from Gibco

High quality sequence stop: 385.

FEATURES
source

1..522
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:315391"
/clone_lib="NCI_CGAP_Man2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 116 a 133 c 158 g 115 t

ORIGIN

Query Match 1.2%; Score 35; DB 10; Length 522;

Best Local Similarity 100.0%; Pred. No. 2.3e-05;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1942 GTCTGTCTCAAGCAAGTCATCAGGCTGGCTG 1976

Db 73 GTCTGTCTCAAGCAAGTCATCAGGCTGGCTG 107

RESULT 7

BF320986

LOCUS uz57h10.y1 NCI_CGAP_Man6 Mus musculus CDNA clone IMAGE:3673219 5' similar to TR:Q9UP80 Q9UP80 METH1 PROTEIN. ;, mRNA sequence.

ACCESSION BF320986

VERSION EST.

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 305)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov/image/html/iresources.shtml

MGI:1433987

High quality sequence stop: 192.

Location/Qualifiers

1..305

FEATURES

source

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3673219"

/clone_lib="NCI_CGAP_Man6"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 79 a 71 c 85 g 69 t

ORIGIN

Query Match 1.1%; Score 30; DB 12; Length 305;

Best Local Similarity 100.0%; Pred. NO. 0.005;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1765 GGCTACACACAGCAGCCACCGGCTCACT 1794

Db 48 GGCTACACACAGCAGCCACCGGCTCACT 77

RESULT 8

AQ440250

LOCUS

DEFINITION

HS_5071.B2.D07.T7A RPCI-11 Human Male BAC Library Homo sapiens

genomic clone Plate=647 Col=14 Row=H, DNA sequence.

ACCESSION AQ440250

VERSION AQ440250.1

KEYWORDS GI:4551589

SOURCE GSS.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 400)

REFERENCE

AUTHORS

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

TITLE

Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

99380589

CONTACT: Mahairas GG, Wallace JC, Hood L

High throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallaceu.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

or from Research Genetics (info@resgen.com). BAC end Web Server:

<http://www.htsc.washington.edu>

Plate: 647 row: H column: 14

Seq primer: T7

Class: BAC ends

High quality sequence stop: 400.

Location/Qualifiers

1..400

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate=647 Col=14 Row=H"

/clone_lib="RPCI-11 Human Male BAC Library"

/sex="male"

/note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor

and partially digested with a combination of EcoRI and

EcoRI Methylase. Size selected DNA was cloned into the

pBAC3.6 vector at EcoRI sites"

BASE COUNT 100 a 116 c 98 g 83 t

3 others

ORIGIN

Query Match 1.0%; Score 29; DB 17; Length 400;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2164 GATGACAACTACTGCTGCTGAAGACAG 2192

Db 359 GATGACAACTACTGCTGCTGAAGACAG 387

RESULT 9

AW485848 415 bp mRNA linear EST 09-JUL-2000
LOCUS 68315 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
DEFINITION AW485848
ACCESSION AW485848
VERSION AW485848.1 GI:7055954
KEYWORDS EST.
SOURCE Pig.

ORGANISM

Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE

AUTHORS Fahnenkrug, S.C., Frenking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
and Keele, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine

JOURNAL

COMMENT Unpublished (2000)

Contact: Smith RPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smithemail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTGCTACGAGG

Plate: 28 row: D column: 18

Seq primer: ATTTAGTGACACTATAG.

Location/Qualifiers

1..415

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone_lib="MARC 1PIG"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT 70 a 141 c 141 g 63 t

ORIGIN

Query Match 1.0%; Score 28; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2808 CCGCAAGCCCCAGGAGCTGGACTTCTGC 2835

Db 35 CCGCAAGCCCCAGGAGCTGGACTTCTGC 62

RESULT 10

B01986 169 bp DNA linear GSS 13-JUL-1996
LOCUS cSRL-145A3-u cSRL flow sorted Chromosome 11 specific cosmid Homo
DEFINITION B01986
sapiens genomic clone cSRL-145A3, DNA sequence.

ACCESSION

B01986

VERSION

B01986.1 GI:1411264

KEYWORDS

GSS.

SOURCE

ORGANISM human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 169)

AUTHORS Evans, G.A., Burbee, D., Davies, C., Hahner, L., Oliver, T., Gilbert, M.,

Jones, D., Ward, T., Gillilan, E., Schagemann, J., Probst, S., Harris

, J., DeFord, J., McFarland, J., Burzinski, K., Khan, M., Kupfer, K. and

Garner, H.R.

Genomic Sequence Sampled Map of Chromosome 11

Unpublished (1996)

Contact: Evans GA, Shane Probst

McDermott Center for Human Growth and Development

University of Texas Southwestern Medical Center At Dallas

5323 Harry Hines Blvd, Dallas TX 75235-8591

Tel: 214-648-1600

Fax: 214-648-1666

Email: gevanseutsw.swmed.edu, shane@mcdermott.swmed.edu

Seq primer: T7

Class: cosmid ends

High quality sequence stop: 169.

Location/Qualifiers

1..169

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="cSRL-145A3"

/clone_lib="cSRL flow sorted Chromosome 11 specific

cosmid"

/sex="female"

/cell_type="chimeric hamster somatic cell hybrid"

/notes="Vector: sCOS-1; Human Chromosome 11 specific cosmid

library prepared from flow sorted human Chromosome 11

derived from Chinese hamster Ovary (CHO) monochromosomal

somatic cell hybrid, J1"

BASE COUNT 26 a 56 c 41 g 36 t 10 others

ORIGIN

Query Match 0.9%; Score 27; DB 17; Length 169;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CGCGGGGTGTTCGGGGCGGCGCTTC 501

Db 56 CGCGGGGTGTTCGGGGCGGCGCTTC 82

RESULT 11

BE281680

LOCUS

DEFINITION

BE281680

ACCESSION

BE281680

VERSION

BE281680.1 GI:9156727

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 627)

AUTHORS NIH-MGC

TITLE NIH-MGC

JOURNAL

COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM8537 row: a column: 16

High quality sequence stop: 530.

4

/note="Organ: kidney; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 193 a 250 c 250 g 156 t
ORIGIN

Query Match 0.8%; Score 25; DB 13; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1700 TGGAGCCCTGCCCGCAGCTCAGCCTC 1724
|||||
Db 540 TGGAGCCCTGCCCGCAGCTCAGCCTC 564

RESULT 14
BE666088
LOCUS 149119 MARC 4BOV Bos taurus cDNA 5', mRNA linear EST 25-APR-2001
DEFINITION BE666088
ACCESSION BE666088
VERSION BE666088.1 GI:10026679
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 464)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: Smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCCGCAGTCACGACG
Plate: 62 row: B column: 2
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1. .464
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from day 20 and day 40 embryos."

BASE COUNT 73 a 157 c 163 g 71 t
ORIGIN

Query Match 0.8%; Score 24; DB 10; Length 464;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2812 AAGCCCCAGGAGCTGGACTTCTGC 2835
|||||
Db 215 AAGCCCCAGGAGCTGGACTTCTGC 238

RESULT 15
BI692986
LOCUS 603344521F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5372261 5', mRNA sequence.
DEFINITION BI692986
ACCESSION BI692986
VERSION BI692986.1 GI:15655615
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 664)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1947 row: j column: 06
High quality sequence stop: 664.
Location/Qualifiers
1. .664
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 133 a 168 c 207 g 156 t
ORIGIN

Query Match 0.8%; Score 24; DB 13; Length 664;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2752 CTCAGTGTGTGGCCACGAGGC 2775
|||||
Db 46 CTCAGTGTGTGGCCACGAGGC 69

Search completed: May 16, 2003, 04:02:41
Job time : 3610 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 22:53:01 ; Search time 340 Seconds
(without alignments)
10824.099 Million cell updates/sec

Title: US-09-965-631-3
Perfect score: 2853
Sequence: 1 atgtcttctgtggcctcct.....gcgtcttgagccgtgtcta 2853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644969091 residues
Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2853	100.0	2853	10	US-09-965-631-3
2	2853	100.0	3446	10	US-09-965-631-7
3	2298	80.5	2469	9	US-10-163-316-3
4	2298	80.5	2940	9	US-10-163-316-1
5	1091.2	38.2	1104	10	US-09-965-631-5
6	959	33.6	966	10	US-09-965-631-1
7	634.4	22.2	2184	9	US-10-097-597-13
8	634.4	22.2	2184	9	US-10-097-580-13
9	634.4	22.2	2184	10	US-09-445-0238-13
10	627.2	22.0	2184	9	US-10-097-597-2
11	627.2	22.0	2184	9	US-10-097-580-2
12	627.2	22.0	2184	10	US-09-445-023A-2
13	626.6	22.0	4676	12	US-10-105-929-1
14	590.4	20.7	3638	10	US-09-918-171A-8
15	581.6	20.4	4407	9	US-10-174-590-351
16	581.6	20.4	4407	9	US-10-176-758-351
17	581.6	20.4	4407	9	US-10-175-737-351
18	581.6	20.4	4407	9	US-10-173-706-351
19	581.6	20.4	4407	9	US-10-175-738-351

20	581.6	20.4	4407	9	US-10-175-752-351	Sequence 351, App
21	581.6	20.4	4407	9	US-10-176-482-351	Sequence 351, App
22	581.6	20.4	4407	9	US-10-176-757-351	Sequence 351, App
23	581.6	20.4	4407	9	US-10-176-913-351	Sequence 351, App
24	581.6	20.4	4407	9	US-10-180-552-351	Sequence 351, App
25	581.6	20.4	4407	9	US-10-180-557-351	Sequence 351, App
26	581.6	20.4	4407	9	US-10-173-700-351	Sequence 351, App
27	581.6	20.4	4407	9	US-10-174-572-351	Sequence 351, App
28	581.6	20.4	4407	9	US-10-174-579-351	Sequence 351, App
29	581.6	20.4	4407	9	US-10-174-582-351	Sequence 351, App
30	581.6	20.4	4407	9	US-10-174-588-351	Sequence 351, App
31	581.6	20.4	4407	9	US-10-175-739-351	Sequence 351, App
32	581.6	20.4	4407	9	US-10-175-740-351	Sequence 351, App
33	581.6	20.4	4407	9	US-10-175-743-351	Sequence 351, App
34	581.6	20.4	4407	9	US-10-176-488-351	Sequence 351, App
35	581.6	20.4	4407	9	US-10-176-492-351	Sequence 351, App
36	581.6	20.4	4407	9	US-10-176-747-351	Sequence 351, App
37	581.6	20.4	4407	9	US-10-176-750-351	Sequence 351, App
38	581.6	20.4	4407	9	US-10-176-985-351	Sequence 351, App
39	581.6	20.4	4407	9	US-10-176-987-351	Sequence 351, App
40	581.6	20.4	4407	9	US-10-176-991-351	Sequence 351, App
41	581.6	20.4	4407	9	US-10-176-992-351	Sequence 351, App
42	581.6	20.4	4407	9	US-10-176-993-351	Sequence 351, App
43	581.6	20.4	4407	9	US-10-184-658-351	Sequence 351, App
44	581.6	20.4	4407	9	US-10-173-695-351	Sequence 351, App
45	581.6	20.4	4407	9	US-10-173-697-351	Sequence 351, App

ALIGNMENTS

RESULT 1
US-09-965-631-3
; Sequence 3, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-3

Query Match	100.0%;	Score 2853;	DB 10;	Length 2853;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2853;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGCTTCTGTCGGCATCCTAAACCTTGGCTTCCCGGGGGAACCGCTGAGGCTT	JAG 60	
Db	1	ATGCTTCTGTCGGCATCCTAAACCTTGGCTTCCCGGGGGAACCGCTGAGGCTT	JAG 60	
QY	61	CCAGAGCGGAGGTAGTCGTTCCATCCGACTGGACCCGACATTAACGGCCGCCCTAC	120	
Db	61	CCAGAGCGGAGGTAGTCGTTCCATCCGACTGGACCCGACATTAACGGCCGCCCTAC	120	
QY	121	TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT	180	
Db	121	TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT	180	
QY	181	CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTTGGTCCCGCTTCTCC	240	
Db	181	CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTTGGTCCCGCTTCTCC	240	
QY	241	ACTGAGCATCTGGCGGTCCCGCTCCAGGGGCTCACCGGGGGCTTCTCAGACCTTGGCAGCG	300	

|||||
Db 241 ACTGAGCATCTGGCGTCCCTCCAGGGCTCACCGGGGCTCTTCAGACCTGGACGC 300
QY 301 TGGTCTATTCTGGGACGTGAACGCCGAGCGGACTCGTTCGCTGCTGTGAGCCCTGTGC 360
Db 301 TGGTCTATTCTGGGACGTGAACGCCGAGCGGACTCGTTCGCTGCTGTGAGCCCTGTGC 360
QY 361 GGGGGCTCCGGGAGCCCTTGGCTACCGAGGCGCGGAGTATGTCAATTAGCCCGCTGCC 420
Db 361 GGGGGCTCCGGGAGCCCTTGGCTACCGAGGCGCGGAGTATGTCAATTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGCGCGAGCGCAACAGCCAGGGCGCACACTCTCCAGCGCGG 480
Db 421 AATGCTAGCGCGCGCGCGAGCGCAACAGCCAGGGCGCACACTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGGGGGGCTTCCGAGAGACCCACCTCTCGCTGGGGGTGCGCTCGGCTGG 540
Db 481 GGTGTTCCGGGGGGGCTTCCGAGAGACCCACCTCTCGCTGGGGGTGCGCTCGGCTGG 540
QY 541 AACCCGCCATCTACGGCCCTTGAACCGCTTACAGCCCGCGGCGGCGGCTTCGGGGAG 600
Db 541 AACCCGCCATCTACGGCCCTTGAACCGCTTACAGCCCGCGGCGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGCGAGGCTTGGCGCGCAAGCGCTTTCGCTCTATCCCGCGGTACGTG 660
Db 601 AGTCGTAGCGCGCGAGGCTTGGCGCGCAAGCGCTTTCGCTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGTGTGTCGCGGACGAGTCAATGGTCAAGTTCACCGGGCGGACCTGGAACAT 720
Db 661 GAGACGCTGTGTGTCGCGGACGAGTCAATGGTCAAGTTCACCGGGCGGACCTGGAACAT 720
QY 721 TATCTGCTGACGCTGTGTCGACGGCGGCGGACTACCGGCATCCGACGATCCTCAAC 780
Db 721 TATCTGCTGACGCTGTGTCGACGGCGGCGGACTACCGGCATCCGACGATCCTCAAC 780
QY 781 CCCATCAACATCTGTGTGTCAGGTGCTCTTAGAGATCGTGACTCCGGGCCCAAG 840
Db 781 CCCATCAACATCTGTGTGTCAGGTGCTCTTAGAGATCGTGACTCCGGGCCCAAG 840
QY 841 GTCACCGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTGCGCAGAAAGCTGAAC 900
Db 841 GTCACCGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTGCGCAGAAAGCTGAAC 900
QY 901 AAGTCAGTGCACAGACACCCGACTGTGGGACTGCCATCCTCTTACAGGACGAGAC 960
Db 901 AAGTCAGTGCACAGACACCCGACTGTGGGACTGCCATCCTCTTACAGGACGAGAC 960
QY 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGCGATGGCTGTGTGGGTACCATGTGAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGACACCCCTGGCGATGGCTGTGTGGGTACCATGTGAC 1020
QY 1021 CCCAAGAGAAGTGTCTGTCAATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1080
Db 1021 CCCAAGAGAAGTGTCTGTCAATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1080
QY 1081 CAGGACTGGCCACGTGTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTG 1140
Db 1081 CAGGACTGGCCACGTGTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCCAGCTGCC 1200
Db 1141 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTCATCCAGATCCAGCTGCC 1200
QY 1201 AACCCCTGTGACGCTGTGCTGCTATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1260
Db 1201 AACCCCTGTGACGCTGTGCTGCTATGAGGACGATGGCTTCCATCAGCCCTCACCACTGCC 1260
QY 1261 GACTGCTCTGAGCAACCCAGGACCCATCTCCCTGCCGAGGATCTGCCGGGCGCC 1320
Db 1261 GACTGCTCTGAGCAACCCAGGACCCATCTCCCTGCCGAGGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCTGTAGCAGCAGTGCAGCTGGCTTTTGGCGTGGCTTCCAAAGCCCTGTCT 1380
|||||

Db 1321 AGCTACACCTGTAGCAGCAGTGCAGCTGGCTTTTGGCGTGGCTTCCAAAGCCCTGTCT 1380
QY 1381 TACATGCAGTACTGCACCAAGCTGTGTGACCGGGAAAGCCCAAGGACAGATGGTGTGC 1440
Db 1381 TACATGCAGTACTGCACCAAGCTGTGTGACCGGGAAAGCCCAAGGACAGATGGTGTGC 1440
QY 1441 CAGACCCGCCACTTCCCTTGGCCGATGGCACCAAGCTGTGGCGAGGGCAAGCTTGCCTC 1500
Db 1441 CAGACCCGCCACTTCCCTTGGCCGATGGCACCAAGCTGTGGCGAGGGCAAGCTTGCCTC 1500
QY 1501 AAAGGGGCTCGTGGAGAGACACAACCTCAACAAGCACAGGGTGGATGCTTCTGGGCC 1560
Db 1501 AAAGGGGCTCGTGGAGAGACACAACCTCAACAAGCACAGGGTGGATGCTTCTGGGCC 1560
QY 1561 AAATGGGATCCTATGCCCCCTGCTCGCGACATGTGGTGGGGCGTGCAGCTGGTAAAG 1620
Db 1561 AAATGGGATCCTATGCCCCCTGCTCGCGACATGTGGTGGGGCGTGCAGCTGGCCAGG 1620
QY 1621 AGGCAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1680
Db 1621 AGGCAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1680
QY 1681 AAATACCGATCTGCAATCTGGAGCCTGCCCCAGCTCAGCCTCCGGAAAGAGTTCGGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCTGCCCCAGCTCAGCCTCCGGAAAGAGTTCGGG 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACACACAGCAACCGGCTCACTCTGCC 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACACACAGCAACCGGCTCACTCTGCC 1800
QY 1801 GTGCATGGGTGCCAAGTACTCGGGCGTGTCTCCCGGACAAAGTGCACAGTCACTCTGC 1860
Db 1801 GTGCATGGGTGCCAAGTACTCGGGCGTGTCTCCCGGACAAAGTGCACAGTCACTCTGC 1860
QY 1861 CGAGCCAAATGCACTGCTACTTCTATGTGCTGSCACCCAAAGTGTGGAGCGCACGCTG 1920
Db 1861 CGAGCCAAATGCACTGCTACTTCTATGTGCTGSCACCCAAAGTGTGGAGCGCACGCTG 1920
QY 1921 TGCTCTCTGACTCCACTCCGCTGTGTCTCAAGGCAAGTGCATCAAGGCTGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACTCCGCTGTGTCTCAAGGCAAGTGCATCAAGGCTGCTGTGAT 1980
QY 1981 GGGAACTGGGCTCCAGAAGAGATTGCAAGTGTGGGCTGTGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGGCTCCAGAAGAGATTGCAAGTGTGGGCTGTGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGTGTGACTGTGACTTTCACCAAGCCCATGATGGCTACAAATTTTCGGTG 2100
Db 2041 AGCTGCAAGAAGTGTGACTGTGACTTTCACCAAGCCCATGATGGCTACAAATTTTCGGTG 2100
QY 2101 GCCATCCCGGAGGCGCTCAAGCATCGACATCGCCAGCGCGGTTCACAAAGGCTGATC 2160
Db 2101 GCCATCCCGGAGGCGCTCAAGCATCGACATCGCCAGCGCGGTTCACAAAGGCTGATC 2160
QY 2161 GGGATCACAACTACCTGGCTCTGAAGAAGCAAGCAAGTACCTGCTCAAGGGCAT 2220
Db 2161 GGGATCACAACTACCTGGCTCTGAAGAAGCAAGCAAGTACCTGCTCAAGGGCAT 2220
QY 2221 TTCGTGTGTGCGGGTGGAGCGGACCTGCTGTAAGGGCAGTCTGCTGCGGTACAGC 2280
Db 2221 TTCGTGTGTGCGGGTGGAGCGGACCTGCTGTAAGGGCAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCACGGGCACAGGGTGGAGAGCCTCGAGCTTCCGGGCCATCTCGAGCGCGTGAAC 2340
Db 2281 GGCACGGGCACAGGGTGGAGAGCCTCGAGCTTCCGGGCCATCTCGAGCGCGTGAAC 2340
QY 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCGGCTCCGCTACTCTCTTCTATCTG 2400
Db 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCGGCTCCGCTACTCTCTTCTATCTG 2400
QY 2401 CCCAAAGAGCCTCGGGAGGACAAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTG 2460
Db 2401 CCCAAAGAGCCTCGGGAGGACAAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTG 2460

QY 2461 TTGCACACAGCGTCTCAGCCTCTCCAAACAGGTGAGCAGCGCGAGCCGCCCT 2520
Db 2461 TTGCACACAGCGTCTCAGCCTCTCCAAACAGGTGAGCAGCGCGAGCCGCCCT 2520
QY 2521 GCACCTGGGTGGTGGCAGCTGGGGCCGTGCTCCGGGAGCTGGCGAGTGGCTGCAG 2580
Db 2521 GCACCTGGGTGGTGGCAGCTGGGGCCGTGCTCCGGGAGCTGGCGAGTGGCTGCAG 2580
QY 2581 AAGCGGGCGGTGAGCTGTCGGGGTTCGCGGGGAGCGACGTCCTGCTGCTGATGCA 2640
Db 2581 AAGCGGGCGGTGAGCTGTCGGGGTTCGCGGGGAGCGACGTCCTGCTGCTGATGCA 2640
QY 2641 GCCCATCGCCCGTGGAGACACAGCCTGCGGGGAGCGCTCCCGACCTGGAGCTCAGC 2700
Db 2641 GCCCATCGCCCGTGGAGACACAGCCTGCGGGGAGCGCTCCCGACCTGGAGCTCAGC 2700
QY 2701 GCCTGGTCACTGCTCCAAAGCTGCGGGCGGGGATTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGGTCACTGCTCCAAAGCTGCGGGCGGGGATTCAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCCACGAGCGCGCTGCTGGCCGGGACCACTGCACTTGCACCGCAAGCCCGAG 2820
Db 2761 GTGGGCCACGAGCGCGCTGCTGGCCGGGACCACTGCACTTGCACCGCAAGCCCGAG 2820
QY 2821 GAGCTGGACTTCTGCTGCTCCTGAGGCCGTGCTGA 2853
Db 2821 GAGCTGGACTTCTGCTGCTCCTGAGGCCGTGCTGA 2853

RESULT 2

US-09-965-631-7
; Sequence 7, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; FILE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Encodin
; TITLE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3446
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-7

Query Match 100.0%; Score 2853; DB 10; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTCTGGCATCCTAACCTGGCTTTTCGGCGGCAACCGCTGGAGGCTCTGAG 60
Db 397 ATGCTTCTCTGGCATCCTAACCTGGCTTTTCGGCGGCAACCGCTGGAGGCTCTGAG 456
QY 61 CCAGACGGGAGGTAGTCTGCTCCATCCGACTGGACCGCGGACATTAACGGCGCGCGCTAC 120
Db 457 CCAGACGGGAGGTAGTCTGCTCCATCCGACTGGACCGCGGACATTAACGGCGCGCGCTAC 516
QY 121 TACTGCGGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACACATTT 180
Db 517 TACTGCGGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACACATTT 576
QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTCTCC 240
Db 577 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTCTCC 636
QY 241 ACTGAGCATCTGGGGGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGACCTGGAGCG 300

Db 637 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTCACACCTGCTCACGC 696
QY 301 TGCTTCTATTTCTGGGACGTGAACGCCGAGCGGACTCTGCTGCTCTGTGAGCCCTGTGC 360
Db 697 TGCTTCTATTTCTGGGACGTGAACGCCGAGCGGACTCTGCTGCTCTGTGAGCCCTGTGC 756
QY 361 GGGGGCTCCGGGAGCCTTTGGCTACCGAGGCGCGAGTATGCTATTAAGCCGCTTGCCC 420
Db 757 GGGGGCTCCGGGAGCCTTTGGCTACCGAGGCGCGAGTATGCTATTAAGCCGCTTGCCC 816
QY 421 AATGCTAGCGCGCGCGGCGAGCGACACAGCAGCGGCGACACCTTCTCCAGCCCGG 480
Db 817 AATGCTAGCGCGCGCGGCGAGCGACACAGCAGCGGCGACACCTTCTCCAGCCCGG 876
QY 481 GGTGTTCCCGGGCGGCTTCCGGAGACCCACCTCTCGCTGCGGGGTGCGCTCGGCTG 540
Db 877 GGTGTTCCCGGGCGGCTTCCGGAGACCCACCTCTCGCTGCGGGGTGCGCTCGGCTG 936
QY 541 AACCCGCCATCTACGGGCCCTTGGACCTTACAAGCGCGGGCGGGGCTTCGCGGAG 600
Db 937 AACCCGCCATCTACGGGCCCTTGGACCTTACAAGCGCGGGCGGGGCTTCGCGGAG 996
QY 601 AGTCTAGCGCGCGAGTCTGGCGCGCAAGCGTTTCTGCTATCCCGCGGTACGTG 660
Db 997 AGTCTAGCGCGCGAGTCTGGCGCGCAAGCGTTTCTGCTATCCCGCGGTACGTG 1056
QY 661 GAGACGCTGCTGTCGCGGACGAGTCAATGCTCAAGTTCACCGCGCGGACCTTGAACAT 720
Db 1057 GAGACGCTGCTGTCGCGGACGAGTCAATGCTCAAGTTCACCGCGCGGACCTTGAACAT 1116
QY 721 TATCTGCTGAGCTGCTGGCAACGGGGCGGACTCTACCGCCATCCAGCATCTCTCAAC 780
Db 1117 TATCTGCTGAGCTGCTGGCAACGGGGCGGACTCTACCGCCATCCAGCATCTCTCAAC 1176
QY 781 CCATCAACATCGTTGGTCAAGTCTGCTCTTAGAGATGCTGACTCCGGCGGCAAG 840
Db 1177 CCATCAACATCGTTGGTCAAGTCTGCTCTTAGAGATGCTGACTCCGGCGGCAAG 1236
QY 841 GTCACCGGCAATCGGCGCTGACGCTCGCAACTTCTGTCCTGGCAGAGAGAGCTGAAC 900
Db 1237 GTCACCGGCAATCGGCGCTGACGCTCGCAACTTCTGTCCTGGCAGAGAGAGCTGAAC 1296
QY 901 AAGTGAAGTCAACAGCACCGCGGACTGAGGACTGCTCTCTTCCACAGCAGGAC 960
Db 1297 AAGTGAAGTCAACAGCACCGCGGACTGAGGACTGCTCTCTTCCACAGCAGGAC 1356
QY 961 CTGTGTGAGCCACCACTGTGACACCTGCGGATGCTGATGGGTACCATGTGTGAC 1020
Db 1357 CTGTGTGAGCCACCACTGTGACACCTGCGGATGCTGATGGGTACCATGTGTGAC 1416
QY 1021 CCCAAGAGAGCTGCTCTGCTCAATTGAGGAGATGGGCTTCCATGAGCTTCCACATTCGC 1080
Db 1417 CCCAAGAGAGCTGCTCTGCTCAATTGAGGAGATGGGCTTCCATGAGCTTCCACATTCGC 1476
QY 1081 CACGAGTGGCGCAGCTGTTCAACATGCCCATCACAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1477 CACGAGTGGCGCAGCTGTTCAACATGCCCATCACAATGTGAAGTCTGTGAGGAGGTG 1536
QY 1141 TTTGGAGAGTCCCGAGCAACCAATGATGTCGCCGACCTCATCCAGATCGACCGTGC 1200
Db 1537 TTTGGAGAGTCCCGAGCAACCAATGATGTCGCCGACCTCATCCAGATCGACCGTGC 1596
QY 1201 AACCCCTGGTCAAGCTGCTGCTGCCATCATCAGGACTTCTGAGAGCGGCGAGT 1260
Db 1597 AACCCCTGGTCAAGCTGCTGCTGCCATCATCAGGACTTCTGAGAGCGGCGAGT 1656
QY 1261 GACTGCTCTGAGCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
Db 1657 GACTGCTCTGAGCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1716
QY 1321 AGCTACACCTTGAGCGAGCTGCGGCTTTTGGCGTGGGCTCCAGCCCTGTCTCT 1380
Db 1717 AGCTACACCTTGAGCGAGCTGCGGCTTTTGGCGTGGGCTCCAGCCCTGTCTCT 1776

```
QY 1381 TACATGCACTACTGCACCAAGCTGTGGTGCACCGGGAGGCCAAGGACACAGATGGTGTGC 1440
DB 1777 TACATGCACTACTGCACCAAGCTGTGGTGCACCGGGAGGCCAAGGACACAGATGGTGTGC 1836
QY 1441 CAGACCCGCCACTTCCCTCGTGGCGGAGTGGACACAGCTGTGGGAGGCCAAGCTGTGCCTC 1500
DB 1837 CAGACCCGCCACTTCCCTCGTGGCGGAGTGGACACAGCTGTGGGAGGCCAAGCTGTGCCTC 1896
QY 1501 AAAGGGGCTCGTGGAGAGACAAACCTCAACAAGCACACAGGCTGGATGCTTCTTGGGCC 1560
DB 1897 AAAGGGGCTCGTGGAGAGACAAACCTCAACAAGCACACAGGCTGGATGCTTCTTGGGCC 1956
QY 1561 AAATGGGATCCTATGATGCGCCCTGTCTCGGCACATGTGGTGGGCGGTGCAGCTGGCCAGG 1620
DB 1957 AAATGGGATCCTATGATGCGCCCTGTCTCGGCACATGTGGTGGGCGGTGCAGCTGGCCAGG 2016
QY 1621 AGGCAGTGCACCAACCCCTGCGCAACGGGGCAAGTACTGCGAGGGAGTGAAGGGTG 1680
DB 2017 AGGCAGTGCACCAACCCCTGCGCAACGGGGCAAGTACTGCGAGGGAGTGAAGGGTG 2076
QY 1681 AAATACCGATCCTGCAATCTGGAGCCCTGCCCAAGCTCAGCCTCCGAAAGAGCTTCCGG 1740
DB 2077 AAATACCGATCCTGCAATCTGGAGCCCTGCCCAAGCTCAGCCTCCGAAAGAGCTTCCGG 2136
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTACAACACAGCACCAACCGGCTCACTCTCGCC 1800
DB 2137 GAGGAGCAGTGTGAGGCTTTCAACGGCTACAACACAGCACCAACCGGCTCACTCTCGCC 2196
QY 1801 GTGCATGGGTGCCCAAGTACTCGGGCTGTCTCCCGGGACAAAGTGCAGGCTCATCTGC 1860
DB 2197 GTGCATGGGTGCCCAAGTACTCGGGCTGTCTCCCGGGACAAAGTGCAGGCTCATCTGC 2256
QY 1861 CGACCAATGGCACTGCTACTTCTATGTGCTGGCACCACCAAGTGTGGACGCGACGCTG 1920
DB 2257 CGACCAATGGCACTGCTACTTCTATGTGCTGGCACCACCAAGTGTGGACGCGACGCTG 2316
QY 1921 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAAGCAAGTGCATTAAGCTGGCTGTGAT 1980
DB 2317 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAAGCAAGTGCATTAAGCTGGCTGTGAT 2376
QY 1981 GGGAACTGGGCTCCAAAGAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAATAAG 2040
DB 2377 GGGAACTGGGCTCCAAAGAGATTCGACAAGTGTGGGTGTGTGGGGAGACAAATAAG 2436
QY 2041 AGCTGCAAGAGTGTGACTGTGACTTTCACCAAGCCCATGATGGCTTACAAATTCGTGGTG 2100
DB 2437 AGCTGCAAGAGTGTGACTGTGACTTTCACCAAGCCCATGATGGCTTACAAATTCGTGGTG 2496
QY 2101 GCCATCCCGCAGCGGCTCAAGCATCGACATCCGCGCGGTTTACAAAGGCTGATC 2160
DB 2497 GCCATCCCGCAGCGGCTCAAGCATCGACATCCGCGCGGTTTACAAAGGCTGATC 2556
QY 2161 GGGGATGACAACCTGCTGCTGAAGAAGCAAGCAAGCAAGTACCTGCTCAACGGCAT 2220
DB 2557 GGGGATGACAACCTGCTGCTGAAGAAGCAAGCAAGTACCTGCTCAACGGCAT 2616
QY 2221 TTGCTGTGTGCGGCTGGAGCGGGACCTGGTGGTGAAGGGCAATCTGCTGGGTTACAGC 2280
DB 2617 TTGCTGTGTGCGGCTGGAGCGGGACCTGGTGGTGAAGGGCAATCTGCTGGGTTACAGC 2676
QY 2281 GGCACGGGCACAGCGTGGAGAGCCCTGACAGGCTTCCCGGCCCATCTGGAGCGGCTGACC 2340
DB 2677 GGCACGGGCACAGCGTGGAGAGCCCTGACAGGCTTCCCGGCCCATCTGGAGCGGCTGACC 2736
QY 2341 GTGAGGCTCTTCTGCTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
DB 2737 GTGAGGCTCTTCTGCTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2796
QY 2401 CCCAAAGAGCCTGGGAGGACAGTCTCTCATCCCAAGGACCCCGGGACCCCTCTGTC 2460
DB 2797 CCCAAAGAGCCTGGGAGGACAGTCTCTCTCATCCCAAGGACCCCGGGACCCCTCTGTC 2856
```

```
QY 2461 TTGCACAACAGCGTCTCTAGCCCTCTCCAAACAGGTGGAGCGGACGACGAGCCGCCCCCT 2520
DB 2857 TTGCACAACAGCGTCTCTAGCCCTCTCCAAACAGGTGGAGCGGACGACGAGCCGCCCCCT 2916
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGGCGGTGCTCCGGAGAGTGGGAGTGGCCCTGCAG 2580
DB 2917 GCACGCTGGGTGGCTGGCAGCTGGGGGCGGTGCTCCGGAGAGTGGGAGTGGCCCTGCAG 2976
QY 2581 AAGCGGGCGGTGACTGTCTGGGGCTCCCGCGGACGACGAGTCCCTGCTGTGATGCA 2640
DB 2977 AAGCGGGCGGTGACTGTCTGGGGCTCCCGCGGACGACGAGTCCCTGCTGTGATGCA 3036
QY 2641 GCCATCGGCCGCTGGAGACAAAGCTGCGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
DB 3037 GCCATCGGCCGCTGGAGACAAAGCTGCGGGGAGCCCTGCCACCTGGGAGCTCAGC 3096
QY 2701 GCCTGGTCACTCTGCTCAAGAGCTGCGGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
DB 3097 GCCTGGTCACTCTGCTCAAGAGCTGCGGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 3156
QY 2761 GTGGGCCACGAGCGCGCTGCTGGCCGGGACCAAGTGCACCTTGCACCGCAAGCCCGAG 2820
DB 3157 GTGGGCCACGAGCGCGCTGCTGGCCGGGACCAAGTGCACCTTGCACCGCAAGCCCGAG 3216
QY 2821 GAGCTGGACTTCTGCTGCTCTGAGGCCGTGCTGA 2853
DB 3217 GAGCTGGACTTCTGCTGCTCTGAGGCCGTGCTGA 3249
```

RESULT 3

```
US-10-163-316-3
; Sequence 3, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP101-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2469)
US-10-163-316-3
```

Query Match

```
Best Local Similarity 80.5%; Score 2298; DB 9; Length 2469;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;
QY 1 ATGCTTCTGCTGGCATCTTAACCCCTGGCTTTCGCCGGGGAACCGCTGGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGCATCTTAACCCCTGGCTTTCGCCGGGGAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGGAGGTAGTCTGCCATCCGACTGGACCCCGACATTAACGGCCGCGGTAC 120
DB 61 CCAGAGCGGGAGGTAGTCTGCCATCCGACTGGACCCCGACATTAACGGCCGCGGTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT 180
DB 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT 180
QY 181 CAGGAGGACTTTTACCTTACACCTGACCGCGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
DB 181 CAGGAGGACTTTTACCTTACACCTGACCGCGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
```

Qy	241	ACTGAGCATCTGGGCGTCCCCTCCAGGGGCTACCGGGGGCTCTTCAGACCTGGACGC	300
Db	241		
Qy	301	TGCTTCTATTCTGGGAGCTGAACCGCAGACCGGACTCTTCGCTGCTGTGAGCCTGTGC	360
Db	301		
Qy	361	GGGGGCTCCGCGAGCCCTTTGGTACCGAGGCGCGAGTATGTCTATTAGCCCGCTGCCCC	420
Db	361		
Qy	421	AATGCTAGCGCCCGCGGCGAGCGGCAACAGCAGGCGCGACACTTCTCCAGCCGCGG	480
Db	421		
Qy	481	GGTGTTCGGGCGGCGCTTCGCGAGACCCCACTCTCGCTCGGGGTGGCTTCGGGCTGG	540
Db	481		
Qy	541	AACCCCGCCATCCTACGGGCGCTTGACACTTTACAAGCGCGGGCGGGCTTCGGGGAG	600
Db	541		
Qy	601	AGTCGTAGCGGGCGAGGTCTGGGCGCGCAAGCGTTTGTGTCTATCCCGCGGTACGTG	660
Db	601		
Qy	661	GACACCTGCTGTGCGGACAGTCAATGTCAAGTTCACGCGCGGACACTGGAAACAT	720
Db	661		
Qy	721	TATCTGCTGACGTGCTGGCAACGGCGGCGACTCTACCGCCATCCCCAGCATCTCTCAAC	780
Db	721		
Qy	781	CCCATCAACATCGTTTGGTCAAGGTGCTGCTTTAGAGATGCTGACTCCGGGCGCAAG	840
Db	781		
Qy	841	GTACCGGCAATCGGCGCTGACGCTGGCAACTCTGTGCTGGCAGAGAGAGCTGAAC	900
Db	841		
Qy	901	AAAGTGAGTGACAAAGCACCCGAGTACTGGGACACTGCCATCTCTTTCACAGCAGGAC	960
Db	901		
Qy	961	CTGTGTGGAGCCACCCTGTGACACCCCTGGGCATGGCTGATGTGGGTACCATGTGTGAC	1020
Db	961		
Qy	1021	CCCAAGAGAAGCTGTCTGTCTATTGAGAGCATGGGCTTCCATCAGGCTTTCACCCTGCC	1080
Db	1021		
Qy	1081	CAGGAGCTGGGCCACCTGTTCAACATGCCCCATGACAATGTGAAGTCTGTGAGGAGGTG	1140
Db	1081		
Qy	1141	TTTGGGAAGCTCGAGCCCAACCACATGATGTCCCCGACCCCTCATCCAGATCGACCGTGCC	1200
Db	1141		
Qy	1201	AACCCCTGTGACCGTGCAGTGTGCCATCATCACCAGCTTCTCTGGACAGCGGCGACGGT	1260
Db	1201		
Qy	1261	GACTGCTCTCTGACCAACCCAGCAAGCCATCTCCCTGCCCGAGGATCTCCGCGGCGCC	1320
Db	1261		
Qy	1321	AGCTACACCTGAGCCAGCAGTGGCGTGTGGCTTTGGCGTGGGCTCCAAAGCCCTGTCTCT	1380
Db	1321		

	Db	1321	 AGTACACCCTGAGCCAGCAGTGCGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCT	1380
	Qy	1381	TACATGCAGTACTGCCAACCAAGCTGTGGTGCACCCGGGAAGGCCAAGGACAGATGGTGTGC	1440
	Db	1381	TACATGCAGTACTGCACCAAGCTGTGGTGCACCCGGGAAGGCCAAGGACAGATGGTGTGC	1440
	Qy	1441	CAGACCCGCCACTTCCTCTGGCCGATGGCACACAGCTGTGGCAGAGGCAAGCTTCACCTC	1500
	Db	1441	CAGACCCGCCACTTCCTCTGGCCGATGGCACACAGCTGTGGCAGAGGCAAGCTTCGCCCTC	1500
	Qy	1501	AAAGGGSCCTGGTGAGAGACACACCTCAACAAAGCACG-----	1541
	Db	1501	AAAGGGSCCTGGTGAGAGACCAACCTCAACAAAGCACAGGCTCCGACTGCATCATTT	1560
	Qy	1542	-----GGTGGATGGT 	1551
	Db	1561	TCCTCAAACAACCTCTATTAAAGGCTACCAATGGCCTGCACACTACACAGGTGGATGGT	1620
	Qy	1552	TCCTGGGCCAATGGGATCCCTATGGCCCTCTCGGCCACATGTGTGGGGCGCTGCAG	1611
	Db	1621	TCCTGGGCCAATGGGATCCCTATGGCCCTCTCGGCCACATGTGTGGGGCGCTGCAG	1680
	Qy	1612	CTGGCAGAGGCGAGTCACCAACCCCACCCCTGCCAACGGGGGCAAGTACTGCGACGGGA	1671
	Db	1681	CTGGCAGAGGCGAGTGCACCAACCCCCCTGCCAACGGGGGCAAGTACTGCGACGGGA	1740
	Qy	1672	GTGAGGTGAATATCCGATCTCGAATCTGGAGCCCTGCCCCAGCTCAGCCTCCGGAAG	1731
	Db	1741	GTGAGGTGAATATCCGATCTCGAATCTGGAGCCCTGCCCCAGCTCAGCCTCCGGAAG	1800
	Qy	1732	AGCTCCGGGAGGACGATGTGAGGCTTTCACGGCTTCAACACAGCACACCAACCCGCTC	1791
	Db	1801	AGCTTCGGGAGGACGATGTGAGGCTTTCACGGCTTCAACACAGCACACCAACCCGCTC	1860
	Qy	1792	ACTCTCGCGTGCATGGTGCCCAAGTACTCCGGCTGTCTCCCGGGAAGTGCAG	1851
	Db	1861	ACTCTCGCGTGCATGGTGCCCAAGTACTCCGGCTGTCTCCCGGGAAGTGCAG	1920
	Qy	1852	CTCATCTCCGAGCCAATGGCATGCTACTTCTATGTCTGGCACCAAGGTGGTGCAC	1911
	Db	1921	CTCATCTCCGAGCCAATGGCATGCTACTTCTATGTCTGGCACCAAGGTGGTGCAC	1980
	Qy	1912	GGCAGCTGTGCTCTCTGACTCCACTCCGTCTGTGTCCAAAGGCAAGTGATCAAGCT	1971
	Db	1981	GGCAGCTGTGCTCTCTGACTCCACTCCGTCTGTGTCCAAAGGCAAGTGATCAAGCT	2040
	Qy	1972	GGCTGTGATGGGAACCTGGGCTCCAAGAGAGATTCGACAGTGTGGGTGTGTGGGGA	2031
	Db	2041	GGCTGTGATGGGAACCTGGGCTCCAAGAGAGATTCGACAGTGTGGGTGTGTGGGGA	2100
	Qy	2032	GACAATAAGAGCTCAAGAAGGTGACTGGACTCTTCACCAAGCCCATCATGGCTACAAT	2091
	Db	2101	GACAATAAGAGCTCAAGAAGGTGACTGGACTCTTCACCAAGCCCATCATGGCTACAAT	2160
	Qy	2092	TTCTGTGTGGCCATCCCGCAGCGCCTCAAGATCGACATCCGCCAGCGGTTCACAA	2151
	Db	2161	TTCTGTGTGGCCATCCCGCAGCGCCTCAAGATCGACATCCGCCAGCGGTTCACAA	2220
	Qy	2152	GGCTGTATCGGGATGACAACCTACCTGGCTCTGAAGACAGCCCAAGTACTGCTC	2211
	Db	2221	GGCTGTATCGGGATGACAACCTACCTGGCTCTGAAGACAGCCCAAGTACTGCTC	2280
	Qy	2212	AACGGGATTTCTGTGTGTGGCGGTGGAGCGGACCTGGTGGTAAGGGCAGTCTGCTG	2271
	Db	2281	AACGGGATTTCTGTGTGTGGCGGTGGAGCGGACCTGGTGGTAAGGGCAGTCTGCTG	2340
	Qy	2272	CGGTACAGCGGCACAGCGGTGGAGCCCTGCAGGCTTCCCGGCCCATCTTGAG	2331
	Db	2341	CGGTACAGCGGCACAGCGGTGGAGCCCTGCAGGCTTCCCGGCCCATCTTGAG	2400
	Qy	2332	CCGTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGCGGTCCG	2395

Db 2401 CGCTGACGTGGAGTCTCTCCGTGGGGAAGATGACACCGGCCCGGACTCTGC 2454

RESULT 4
US-10-163-316-1
; Sequence 1, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MPI01-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIORITY FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2940)
; OTHER INFORMATION: n = A,T,C or G

US-10-163-316-1

Query Match 80.5%; Score 2298; DB 9; Length 2940;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;

Qy 1 ATGCTTCTGCTGGGCATCTTAACCTTTCGCGGGGGAACCGCTGGAGCTCTGAG 60
Db 472 ATGCTTCTGCTGGGCATCTTAACCTTTCGCGGGGGAACCGCTGGAGCTCTGAG 531

Qy 61 CCAGAGCGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGCCGCCCTAC 120
Db 532 CCAGAGCGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGCCGCCCTAC 591

Qy 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCCGGGACTCATTTTCAGATCAGCATTT 180
Db 592 TACTGGCGGGTCCCGAGGACTCCGGGGATCCGGGACTCATTTTCAGATCAGCATTT 651

Qy 181 CAGGAGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTTGGCTCCCGCTTCTCC 240
Db 652 CAGGAGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTTGGCTCCCGCTTCTCC 711

Qy 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCAGGGGCTTCAGACCTGGACGC 300
Db 712 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCAGGGGCTTCAGACCTGGACGC 771

Qy 301 TGCTTCTATTCTGGGACGTGAACCGCGGAGCCGACTCTGCTGCTCTGTAGCCCTGTCG 360
Db 772 TGCTTCTATTCTGGGCGGTGAACCGCGGAGCCGACTCTGCTGCTCTGTAGCCCTGTCG 831

Qy 361 GGGGGGCTCCCGGAGCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCCGCTGCC 420
Db 832 GGGGGGCTCCCGGAGCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCCGCTGCC 891

Qy 421 AATGCTAGCGCGCGGGGGAGCGCAACAGCGGGCGCACCTTCTCCAGGCGCGG 480
Db 892 AATGCTAGCGCGCGGGGGAGCGCAACAGCGGGCGCACCTTCTCCAGGCGCGG 951

Qy 481 GGTGTTCCGGGGGCTTCCGGAGACCCACTCTCTGCTGGGGTGGCTCTGGGCTGG 540
Db 952 GGTGTTCCGGGGGCTTCCGGAGACCCACTCTCTGCTGGGGTGGCTCTGGGCTGG 1011

Qy 541 AACCCGCCATCTCAGGGGCCCTGGACCTTAAAGCCCGCGCGGGGCTTCGGGGAG 600
Db 951 AACCCGCCATCTCAGGGGCCCTGGACCTTAAAGCCCGCGCGGGGCTTCGGGGAG 2151

Db 1012 AACCCGCCATCTCAGGGGCCCTTGACCCCTTAAAGCCCGCGGGGCTTCGGGAG 1071

Qy 601 AGTCGTAGCGCGGAGGTCTGGCGCGCAAGCGCTTTCGTCTCTATCCCGGGTACGTG 560
Db 1072 AGTCGTAGCGCGGAGGTCTGGCGCGCAAGCGCTTTCGTCTCTATCCCGGGTACGTG 1131

Qy 661 GAGACGCTGGTGTGTCGGGACGAGTCAATGGTTCACGGCGCGGACCTGGAACAT 720
Db 1132 GAGACGCTGGTGTGTCGGGACGAGTCAATGGTTCACGGCGCGGACCTGGAACAT 1191

Qy 721 TATCTGCTACGCTGCTGCAACGGCGCGGACTCTACCGCCATCCAGCATCTCAAC 780
Db 1192 TATCTGCTACGCTGCTGCAACGGCGCGGACTCTACCGCCATCCAGCATCTCAAC 1251

Qy 781 CCATCAACATCGTTGTGTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
Db 1252 CCATCAACATCGTTGTGTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCCCAAG 1311

Qy 841 GTACCGGCAATCGCGCCCTGACGCTGGCGCAACTTCTGTGCTGGCAGAAAGCTGAAC 900
Db 1312 GTACCGGCAATCGCGCCCTGACGCTGGCGCAACTTCTGTGCTGGCAGAAAGCTGAAC 1371

Qy 901 AAAGTGAGTGACAAGACCCCGAGTACTGGGACACTGCCATCCTCTTACCAGGCAAC 960
Db 1372 AAAGTGAGTGACAAGACCCCGAGTACTGGGACACTGCCATCCTCTTACCAGGCAAC 1431

Qy 961 CTGTGTGGAGCCACCACCTGTGACACCTGGGCGATGGTGTGGTACCATGTGAC 1020
Db 1432 CTGTGTGGAGCCACCACCTGTGACACCTGGGCGATGGTGTGGTACCATGTGAC 1491

Qy 1021 CCCAAGAGAAGTGTCTGTCTATTGAGGACGATGGGCTTCCATCAGCCTTCCACCATGCC 1080
Db 1492 CCCAAGAGAAGTGTCTGTCTATTGAGGACGATGGGCTTCCATCAGCCTTCCACCATGCC 1551

Qy 1081 CAGGAGCTGGGCGAGCTGTCAACATGCCCATGCCCATGCAATGTGAAGTCTGTGAGAGGTG 1140
Db 1552 CAGGAGCTGGGCGAGCTGTCAACATGCCCATGCCCATGCAATGTGAAGTCTGTGAGAGGTG 1611

Qy 1141 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGGACCTTCATCCAGATCGACTGCC 1200
Db 1612 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGGACCTTCATCCAGATCGACTGCC 1671

Qy 1201 AACCCCTGGTACGCTGAGTGTGCCATCATCACCGACTTCTCTGGACAGCGGCGAGGT 1260
Db 1672 AACCCCTGGTACGCTGAGTGTGCCATCATCACCGACTTCTCTGGACAGCGGCGAGGT 1731

Qy 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCGCGAGGATCTCCGCGGCGCC 1320
Db 1732 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCGCGAGGATCTCCGCGGCGCC 1791

Qy 1321 AGCTACACCTCTGAGCCAGTGTGCGAGTGTGCTGCTGCTGGGCTCCAAAGCTCTTCT 1380
Db 1792 AGCTACACCTCTGAGCCAGTGTGCGAGTGTGCTGCTGCTGGGCTCCAAAGCTCTTCT 1851

Qy 1381 TACATGACGACTTCCACCAAGCTGTGTGACCGGAGGCGCAAGGACAGATGCTGTC 1440
Db 1852 TACATGACGACTTCCACCAAGCTGTGTGACCGGAGGCGCAAGGACAGATGCTGTC 1911

Qy 1441 CAGACCCGCCACTTCCCTGCGCGGATGGCACCAAGCTGTGCGGAGGCAAGCTCTGCTC 1500
Db 1912 CAGACCCGCCACTTCCCTGCGCGGATGGCACCAAGCTGTGCGGAGGCAAGCTCTGCTC 1971

Qy 1501 AAAGGGGCTCGTGGAGAGACACAACCTCAACAGCACAG----- 1541
Db 1972 AAAGGGGCTCGTGGAGAGACACAACCTCAACAGCACAGGCGCTCCGACTGACATCAT 2031

Qy 1542 -----GGTGGATGGT 1551
Db 2032 TCTCCAAAACAACCTCTTATTAAAGGCTACCAATGGCCTGCACACTACACAGGTGGATGT 2091

Qy 1552 TCCTGGGCGAAATGGGATCCCTATGGCCCTGTCTCGGCGACATGTGTGGGGCGGTGAC 1611
Db 2092 TCCTGGGCGAAATGGGATCCCTATGGCCCTGTCTCGGCGACATGTGTGGGGCGGTGAC 2151

QY 1021 CCCAAGAGAGCTGCTGTCTATTGAGGACGATGGCTTCCATCAGCCTTACCACTGCC 1080
|||||
Db 1021 CCCAAGAGAGCTGCTGTCTATTGAGGACGATGGCTTCCATCAGCCTTACCACTGCC 1080
QY 1081 CACGAGCTGGGCCAGG 1096
|||||
Db 1081 CACGAGCTGGGTAAGG 1096

RESULT 6

US-09-965-631-1
; Sequence 1, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965, 631
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236, 689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 966
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-1

Query Match 33.6%; Score 959; DB 10; Length 966;
Best Local Similarity 100.0%; Pred. No. 2.1e-248;
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGCATCTTAACCTTCGGCTTTCGCGGGGAAACCGCTGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGCATCTTAACCTTCGGCTTTCGCGGGGAAACCGCTGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTGTCTCCATCCGACTGGACCCGACATTAACGCGCCCGCTAC 120
Db 61 CCAGAGCGGAGGTAGTGTCTCCATCCGACTGGACCCGACATTAACGCGCCCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCGGGGATCAGGACTCATTTTCAGATCACAGCATTT 180
Db 121 TACTGGCGGGTCCCGAGGACTCGGGGATCAGGACTCATTTTCAGATCACAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGCTTTCAGACCTTGGACGC 300
Db 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGCTTTCAGACCTTGGACGC 300
QY 301 TGCTTCTATTCTGGGAGCTGAACCGGACCGGAGCTCGTTCGCTGCTGAGCCTGTGC 360
Db 301 TGCTTCTATTCTGGGAGCTGAACCGGAGCGGAGCTCGTTCGCTGCTGAGCCTGTGC 360
QY 361 GGGGGGCTCCCGGAGCTTTTGGCTACCGAGGCGCGGAGTATGTCATTAGCCCGCTGCC 420
Db 361 GGGGGGCTCCCGGAGCTTTTGGCTACCGAGGCGCGGAGTATGTCATTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGGCGAGCGCAACAGCCAGGGCGCACACTTCTCCAGCGCGCG 480
Db 421 AATGCTAGCGCGCGGGGCGAGCGCAACAGCCAGGGCGCACACTTCTCCAGCGCGCG 480
QY 481 GGTGTTCCGGGGGGGCTTTCCGAGACCCACTCTCTCGCTCGGGGTGCTCGGGCTGG 540
Db 481 GGTGTTCCGGGGGGGCTTTCCGAGACCCACTCTCTCGCTCGGGGTGCTCGGGCTGG 540
QY 541 AACCCCGCATCTCAGGGGCCCTGGACCTTACAGCCGCGCGGGGCTTCGGGGAG 600
|||||

Db 541 AACCCCGCATCTCAGGGGCCCTGGACCTTACAGCCGCGCGGGGCTTCGGGAG 600
QY 601 AGTCGTAGCGCGGAGGTCTGGGCGCGCCAAAGCGTTTCGTCTATCCCGGGTACGTG 660
|||||
Db 601 AGTCGTAGCGCGGAGGTCTGGGCGCGCCAAAGCGTTTCGTCTATCCCGGGTACGTG 660
QY 661 GAGACGCTGCTGTCGCGGACGAGTCAATGTCAGATTCCACGGCGCGGACCTGGAACAT 720
Db 661 GAGACGCTGCTGTCGCGGACGAGTCAATGTCAGATTCCACGGCGCGGACCTGGAACAT 720
QY 721 TATCTGCTGACGCTGCTGSCAACGGCGGCGGACTTACCGCCATCCAGCATCTCAAC 780
|||||
Db 721 TATCTGCTGACGCTGCTGSCAACGGCGGCGGACTTACCGCCATCCAGCATCTCAAC 780
QY 781 CCATCAACATCGTTGTCGTCAGGTGCTGCTTTCAGATCGTGACTCGGGCCCAAG 840
Db 781 CCATCAACATCGTTGTCGTCAGGTGCTGCTTTCAGATCGTGACTCGGGCCCAAG 840
QY 841 GTACCGGCAATGCGGCCCTGACGCTGGCGAACTTCTGTGCTGGCAGAAAGCTGAAC 900
Db 841 GTACCGGCAATGCGGCCCTGACGCTGGCGAACTTCTGTGCTGGCAGAAAGCTGAAC 900
QY 901 AAGTGAGTGACAAGACACCCCGAGTACTGGGACACTGCCATCTTTCACCAGGCAGGA 959
Db 901 AAGTGAGTGACAAGACACCCCGAGTACTGGGACACTGCCATCTTTCACCAGGCAGGA 959

RESULT 7

US-10-097-597-13
; Sequence 13, Application US/10097597
; Publication No. US20030022352A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inouchi, Eiji
; APPLICANT: Hakozaki, Michinori
; APPLICANT: Ishioaka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same,
; TITLE OF INVENTION: pharmaceutical
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADA
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/10/097,597
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 2184
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(2184)
US-10-097-597-13

Query Match 22.2%; Score 634.4; DB 9; Length 2184;
Best Local Similarity 60.6%; Pred. No. 5.5e-161;
Matches 1097; Conservative 0; Mismatches 701; Indels 12; Gaps 3;

QY 631 AAGCGTTTCGTGTCTATCCCGGTACGTGAGACGCTGGTGGCGGAGCTCAATG 690
Db 31 AAGCGATTGTGTCCAGCCCGCTTATGTGAAACCAATGCTCGTAGCTGACCATG 90
QY 691 GTCAAGTTCACGCGCGGACCTGGAACATATTCGTGACGCTGTGTCGCAACGGCGCG 750
Db 91 GCCGACTTCCAGCGCAGCGGTCTAAAGCATTTACCTTCTAACCTGTTCCTGGTGGCAGCC 150
QY 751 CGACTCTACGCCCATCCAGCATCTCTCAACCCCATCAACATCGTTGTGGTCAAGGTGCTG 810

151	AGGTTTTACAAGCATCCACGCAATTTAGGAATTTCAATTTAGCCTGGTGTGCTGAGATCTTG	210
811	CTTCTTAGAGATCTGACTCGGGCCCAAGGTACCGGCAATGCGGCCCCIACAGCTGTGGC	870
211	GTCAATACGAGGAGCAGAAGGACCAAGATTACCTCCAATGCAGCTCTCACCTCTGG	270
871	AACCTTCTGTGCTGGCAGAAAGCTGAACAAATGAGTGACAAGCACCCGAGTACTGG	930
271	AATTTCTGCAGCTGSCAGAAACACACACAGCCCCAGTGACCGGATCCAGAGCACTAT	330
931	GACACTGCCATCTCTTCACAGCGAGGACTGTGTGGAGCCACACCTGTGACACCCCTG	990
331	GACACTGCAATTTCTGTTCACACAGACGATTTATGTGGTCTCCACACGCTGACACTCTC	390
991	GGCATGGTGTATGGGTACCATGTGTGACCCCAAGAGAGCTGCTCTCATTCAGGAC	1050
391	GGATGGCAGATGTGGAACCGTATGTACCCACGACGAGCTGCTCAGTCATAGAAGAT	450
1051	GATGGCTTCATCAGCCCTTTCACACTGCCAGAGCTGGCCACGCTGTTCACATGCCCC	1110
451	GATGGTTTGAAGTGGCTTCACACAGCCATGAATTGGGCCATGTGTTTAACATGCG	510
1111	CATGACAATGTGAAGTCTGTGAGGAGTGTGTTGGGAAGCTCCGAGCAACACATGATG	1170
511	CACGATGATGCTAAGCACTGTGCCAGCTTGAATGTGTGAGTGGCGATTTCTCATCTGATG	570
1171	TCCCGACCTCTATCCAGATCGACCTTGCCAAACCCCTGTGCTCAGCTGCAGTGTGCCATC	1230
571	GCCTCGATGCTCTCCAGCTTAGACCATAGCCAGCCCTGCTCATTGCACTGCCTACATG	630
1231	ATCACCCACTTCTTGGACACGGGACGGTGACTGTGCTCTCTGGACCAACCCAGAACGCC	1290
631	GTCACGTCTCTCTAGATAATGACACGGGGAATGTTGATGCAAGCCCGCAGAAATCCA	690
1291	ATCTCCCTGCCCAGGATGTCGCGGCGCCAGCTACACCTCGAGCGACAGCTGCGAGCTG	1350
691	ATCAAGCTCCCTCTGATCTTCCGGTACCTGTGACGATGCCAACGCCAGCTGTACGTTT	750
1351	CTTTTGGGTGGCTCCAAAGCCCTGTCTTTACATGCAGTAG --- TGCACCAAGCTGTGG	1407
751	ACATTCGAGAGGAATCCAAAGCACTGCCCTGATGACGCCAGCACATGTACTACCTCTGG	810
1408	TGCACCGGAAGGCCAAGGACAGATGTGTGCCAGCCGCCACTTCCCTGTGGCCGAT	1467
811	TGCACTGGCACTTCGGTGCCTTACTTGGTGTGCCAAACAAACACTTCCCTTTGGGCAAGT	870
1468	GGCACCAAGCTGTGCGGAGGCAAGCTCTGCCCTCAAAGGGGCTGCTGGTGAGAGACACAAC	1527
871	GGCACCAAGCTGTGAGAAGGGAAGTGTGTGTCAGTGGCAAGTGGTGAACAACAGACAGAC	930
1528	CTCAACAAGCACAG - - - - - GGTGGATGTTCTCTGGGCCAAATGGGATCTCCTATGGCCCC	1581
931	ATGAAGCATTTTGTCTACTCTCTGTTCAATGGAAGCTGGGACCATGGGACCTTGGGGAGAC	990
1582	TGCTCGGCACTATGTGTGGGGCGGTGACAGCTGGCCAGGAGCAAGTGCACCAACCCCAAC	1641
991	TGCTCAAGAACCCTGTGTGTGGAGTTCAATACACAATGAGAGAAATGTGACAAACCCAGTC	1050
1642	CTGCCAACCGGGCAAGTACTGCGAGGAGTGTGGGTGAAATACCGATCTTGCATCTG	1701
1051	CCAAAGAACGGAGGGAAGTACTGTGAAGCAACACAGTCCGCTACAGGTCTCTGTAAACATC	1110
1702	GAGCCCTGCCCAAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAAGTGTGAGGCTTTC	1761
1111	GAGGACTGTCCAGACAATAA --- CGGAATAAGCTTCAGAGAGAGCAAGTGTGCGGCGCAC	1167
1762	AAGGCTTACAACACAGCACCAACCGGCTCACTCTCGCCGTGGCATGGGTGCCCAAGTAC	1821
1168	AATGAGTTTTTCCAAAGCTTCTTTTGGGAATGAGCCCACTGTAGAGTGGACACCCAGTAC	1227
1822	TCCGGGCTGTCTCCCGGACAAAGTGCAGCTCATCTCCCGAGGCAATGGCACTGGCTAC	1881

RESULT 8

RESULT 8
US-10-097-580-13

US-10-097-380-13
: Sequence 13, Application US/10097580; sequence 13, Application US/100
; Publication No. US20030032168A1; PUBLICATION NO. US20
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: Hirose, Kunitaka

APPLICANT: Inoguchi, Ei'ji

APPLICANT: HAKOZAKI, Michinori

APPLICANT: Ishioka, Keiko

; APPLICANT: Ishida, Yukako

APPLICANT: Matsushima,

APPLICANT: Kuno, Kouji

TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical

; TITLE OF INVENTION: composition and meth
 ; FILE REFERENCE: 057002

FILE REFERENCE: Q57092
CURRENT APPLICATION NUMBER: HS10109

; CURRENT APPLICATION NUMBER: US/10/09
 : CURRENT FILING DATE: 2002-03-15

; PRIOR APPLICATION NUMBER: 09/445,023

; PRIOR FILING DATE: 1999-12-03

; PRIOR APPLICATION NUMBER: JP 9-160422

; PRIOR FILING DATE: 1997-06-03

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Pat

; SEQ ID NO 13
: LENGTH: 2184

```
; LENGTH: 2.
: TYPE: DNA
```

TYPE: DNA
ORGANISM: Muis sp

ORGANISM:
FEATURE:

```
; NAME/KEY: exon
```

; LOCATION: (1)..(2184)

US-10-097-580-13

Query Match 22.2%; Score 634.4; DB 9; Length 2184;
Best Local Similarity 60.6%; Pred. No. 5.5e-161;
Matches 1097; Conservative 0; Mismatches 701; Indels 12; Gaps 3;

```
QY 631 AAGGTTTCGTGCTATPCCGCGGTCAGTGGAGACGCTGGTGGTCGCGGACGAGTCAATG 690
DB 31 AAGCGATTGTGTCAGCGCCCGCTTATGTGGAACCACTGCTAGCTGACCAAGTCCATG 90

QY 691 GTCAAGTTCCACGGCGGACCTGGACATTAATCTGCTGACGCTGCTGCAACGGCGG 750
DB 91 GCGGACTTTCACGGCAGCGCTTAAAGCAATTAACCTTCTAACCCTGTCTTGGTGGCAGC 150

QY 751 CGACTCTACCGCCATCCACGATCCTCAACCCCATCAACATCGTTGTGGTCAAGTGTG 810
DB 151 AGGTTTACAGCATCCACGATAGGAATTCATTTAGCCTGGTGGTGAAGATCTTG 210

QY 811 CTCTTTAGAGATCGTACTCCGGGCGCAAGGTCAACGGCAATGCGGCCCTGACGCTGCG 870
DB 211 GTCATATACGAGGAGCAGAAAGGACCAAGAGTTACCTCCAATGACGCTCTCACCCCTCG 270

QY 871 AACTTCTGCTGCGGACAGAGAGCTCAACAAAGTGAGTGACAAGCACCCCGAGTACTGG 930
DB 271 AATTTCTGCACTGGCGAAGAACACACACAGCCCGAGTGACCGGATCCAGAGCACTAT 330

QY 931 GACACTGCCATCTCTTCCACGAGCAGGACCTGTGTGGAGCCACCACTGTGACACCCCTG 990
DB 331 GACACTGCAATCTGTTCACAGACAGGATTTATGTGGCTCCACACAGTGTGACACTCTC 390

QY 991 GGCATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGAGCTGCTGTGATAGGAG 1050
DB 391 GGAATGCGAGATGTTGGAACCGTATGTGACCCAGCAGGAGCTGCTCAGTCAATAGAAT 450

QY 1051 GATGGCTTCCATCAGCTTCCACACTGCCAGAGCTGGGCCAGCTGTTCAACATGCC 1110
DB 451 GATGTTTGAAGTGGCTTCCACAGCCCATGAATTTGGGCCATGTGTTTAACATGCCG 510

QY 1111 CATGACAAATGTGAAGTCTGTGAGGAGTGTGTTGGGAAGCTCCGAGCCCAACCATGATG 1170
DB 511 CACGATGATGCTAAGCACTGTCCAGCTTGAATGGTGTGAGTGGCGATTTTCATCTGATG 570

QY 1171 TCCCGGACCTCATCCAGATGACCGTGGCCAAACCCCTGTGACGCTGCAAGTGGCCATC 1230
DB 571 GCCTCGATGCTCTCCAGCTTAGACCATAGCCAGCCCTGTGTCACCTTGCATGGCTACATG 630

QY 1231 ATCAACGACTTCTGGACAGCGGCGGTGACTGCTCTGTGACCAATCCAGCAAGCCC 1290
DB 631 GTCAGTCTCTTAGATATATGACAGGGGATGTTGTATGACAAAGCCCAAGATCCA 690

QY 1291 ATCTCCCTGCGGAGGATCTGCGGGGCGGAGCTACACCTGAGCCAGAGTGGAGCTG 1350
DB 691 ATCAAGCTCCCTCTGATCTTCCGGTACCTTTGATGATGCCAACCGGCGAGTGCAGTTT 750

QY 1351 GCTTTTGGCTGGCTCCAGGCTGCTGCTTACATGCACTAGTAC---TGCACCAAGCTGTGG 1407
DB 751 ACATTCGGAGAGGAATCCAGACACTGCGCTGATGACGACGACATGCTACTACCTGTGG 810

QY 1408 TGCACCGGAGGCAAGGACAGATGCTGTGCGACAGCCGCACTTCCCTTGGGCGCGAT 1467
DB 811 TGCAGTGGCACTCCGCTGGCTTACTGTGTGCGCAACAAACATCTCCCTTGGGCGAGAT 870

QY 1468 GGCACAGCTGTGGCGAGGCAAGCTGTGCTCAAGGGGCGTGGTGGAGAGACACAC 1527
DB 871 GGCACAGCTGTGGAGAGGGAAGTGTGTGAGTGGCAAGTGCSTGGAACAGAGACAGAC 930

QY 1528 CTCACAGACACAG-----GGTGGATGTTCTGGGCCAAATGGGATCCCTATGGCCCC 1581
DB 931 ATGAAGCATTTTGTACTCTCTGTTTCAATGAGCTGGGAGCCATGGGACCGTGGGAGAC 990

QY 1582 TGCTCGGCGACATGTGGTGGGCGGTGCACTGGCCAGGAGGAGTGCACCAACCCCAACC 1641
DB 991 TGCTCAAGAACCTGTGCTGGTGGAGTTCAATACACATGAGAGAAATGTGACAAACCCGATC 1050
```

```
QY 1642 CCTGCCAACGGGGGCAAGTACTGCGAGGAGTGAAGGTGAATPACCATCTCTCAATCTG 1701
DB 1051 CCAAGAAACGGGAGGAGTACTGTGAAGCAACAGAGTCCGCTACAGTCTCTGTGAATC 1110

QY 1702 GAGCCCTCCCCAGCTCAGCCTCCGGAAGAGCTTCGGGAGGAGCAGTGTGAGGCTTTC 1761
DB 1111 GAGGACTGTCCAGACAATAA---CGGAAAAACGTTTACAGAGGAGCAGTGGAGGCCAC 1167

QY 1762 AACGGCTACAACACAGCAACCAACGCGCTACTCTCGCCGTGGCATGGGTGCCCAAGTAC 1821
DB 1168 AATGATTTTCCAAAGCTTCTTTGGGAATGAGCCCACTGTAGATGACACCCCAAGTAC 1227

QY 1822 TCCGGCGTGTCTCCCGGGACAAGTGAAGCTATCTGCGGAGCCAAATGGCAGCTGCTAC 1881
DB 1228 GCGGCGTCTCGCCAAAGGACAGGTGCAAGCTCACTGTGAAGCCAAAGGCAATTTGGTAC 1287

QY 1882 TTTATGTCGTGGCACCAAGGTGTGACGCGACGCTGTCTCTCTGACTCCACCTCC 1941
DB 1288 TTTTTCGCTTTACAGCCCAAGGTGTAGATGGCACTCCCTGTAGTCCAGACTTACCTCT 1347

QY 1942 GTCTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGATGGGAACCTGGGCTCCAAGAA 2001
DB 1348 GTCTGTGTGCAAGGCACTGTGTGAAGCTGGTGTGATGTCATCATAGACTTCCAAAGAA 1407

QY 2002 AGATTTCGACAGTGTGGGTTGTGGGGGAGACAATAGAGCTGCAAGAGGTGACTGGA 2061
DB 1408 AAGTTTGTAGTGTGGGCTTTGTGGAGGAACCGTTCACATGCAAGAGATGTCAGGA 1467

QY 2062 CTCTTCCACCAAGCCATCATGGCTTACAAATTTCTGTGGTGGCCATCCCGCAGGCGCTCA 2121
DB 1468 ATAGTCACTAGTACAAAGACCTGGGTATCATGACATTTGACAAATCTCTGTGGAGCCACC 1527

QY 2122 AGCATCGACATCCCGCAGCGGTTTACAAAGGCTGTATCGGGGATGACAACTACCTGGCT 2181
DB 1528 RACATTGAAGTGAACATCGGAATCAAAGGGGTCCAGAAACAATGGCAGCTTTCTGGCT 1587

QY 2182 CTGAAGAACAGCCAGGCAAGTACTGCTCAACGGGCAATTTCTGGTGTGGCTGGCGTGGAG 2241
DB 1588 ATTAGAGCGCTGTAGTGTACCTATATTTCTGAATGGAACCTTCACTCTGTCCACACTAGAG 1647

QY 2242 CGGGACCTGGTGTGAAGGCGAGTCTGCTGCGGTACAGCGGCAGCGGACAGCGGTGGAG 2301
DB 1648 CAAGACCTCACTACAAAGGTACTGCTTTAAGGTACAGTGGTCTCGGCTGCGTGGAA 1707

QY 2302 AGCCTGCGAGGCTTCCCGCCCATCTGAGCGCGTGTACCGGTGGAGGTCTCTCGCTGGGG 2361
DB 1708 AGAATCCGCGAGCTTTAGTCCACTCAAAGAACCTTTAACCATCCAGGTTCTTTATGTAGGC 1767

QY 2362 AGATGACACCGCCCGGTCGGTACTCTCTATCTGCCCCAAAGAGCCCTCGGAGGAC 2421
DB 1768 CATGCTCTCCGACCCCAAAATTAATTCACCTACTTTTATGAAGAGACAGACAGTCAATC 1827

QY 2422 AAGTCTCTCTC 2431
DB 1828 AACGCCATTC 1837
```

RESULT 9

US-09-445-023A-13
; Sequence 13, Application US/09445023A
; Patent No. US20020119167A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; composition and method of immunologically analyzing human ADA
; FILE REFERENCE: Q57092

Sequence 2, Application US/10097597

Publication No. US20030022352A1

GENERAL INFORMATION:

APPLICANT: Hirose, Kunitaka

APPLICANT: Inoguchi, Eiichi

APPLICANT: Hakozaaki, Michinori

APPLICANT: Ishioka, Keiko

APPLICANT: Ishida, Yukako

APPLICANT: Matsushima, Kouji

APPLICANT: Kuno, Kouji

TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same,

TITLE OF INVENTION: pharmaceutical

TITLE OF INVENTION: composition and method of immunologically analyzing human ADAMTS

FILE REFERENCE: Q57092

CURRENT APPLICATION NUMBER: US/10/097,597

CURRENT FILING DATE: 2002-03-15

PRIOR APPLICATION NUMBER: 09/445,023

PRIOR FILING DATE: 1999-12-03

PRIOR APPLICATION NUMBER: JP 9-160422

PRIOR FILING DATE: 1997-06-03

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn version 3.0

SEQ ID NO 2

LENGTH: 2184

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: exon

LOCATION: (1)..(2184)

US-10-097-597-2

Query Match 22.0%; Score 627.2; DB 9; Length 2184;

Best Local Similarity 60.8%; Pred. No. 4.8e-159;

Matches 1079; Conservative 0; Mismatches 683; Indels 12; Gaps 3;

```
QY 631 AAGCGTTTCGTCTATCCCGCGTACGTGGAGACCTGGTGGTCCGCGACGAGTCAATG 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 31 AAGCGATTTGTGTCAGCCGCCCTTATGTGGAACCATGCTGTGGCAGACCATGCGATG 90
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 691 GTCAAGTTTCCACGGCGGACCTTGAACATTAATCTCTGACGCTGTGTCGCAACGGCGG 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 91 GCAGAAATCCACGGCAGTGTCTTAAGCATTAACCTTCTCACGTTGTTTTCGGTGGCAG 150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 751 CGACTTACCGCCATCCAGCATCTCAACCCCATCAATCGTTGTTGGTCAAGTGTCTG 810
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 151 AGATTGTACAAACACCCAGCATTCGTAATTCAGTTAGCTGTGTTGTTGTAAGATCTTG 210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 811 CTCTTAGAGATCGTGAATCCGGGCCCAAGTGCACCGCAATCGCGCCCTGACGCTGCGC 870
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 211 GTCAATCCAGATGACAGAGGGCGGAGTGAATGACCTCCATGCTGCCCTCACTCTGGCG 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 871 AACTTCTGTGCTGGCAGAGAGCTGAACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 930
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 271 AACTTTTGAACGTGGCAGAGAGCAGCAACCCACCCAGTGACCGGATGCAGAGCACTAT 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 931 GACACTGCCATCTTTCACCGAGGAGGACCTGTGTGGAGCCACCACTGTGACACCTTG 990
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 331 GACACGAATCTTTTACACAGAGGACTTGTGTGGTCCAGACATGTGTATCTTT 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 991 GGCATGCTGATGTGGTGTACCATGTGTGACCCCAAGAGAGCTGTCTGTTCATTGAGGAC 1050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 GGGATGCTGATGTGGAATGTGTGTGATCCGAGCAGAGCTGCTCGGTATAGAGAT 450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1051 GATGGCTTCCATCAGCTTTCACACTGCGCCAGAGCTGGGCCACGCTGTTTCAACATGCC 1110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 GATGGTTTACAAGCTGCTTCCACAGCCCATGAATAGGCCACGCTGTTTAACATGCCA 510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1111 CATGACAAATGTGAAGTCTGTGAGGAGTGTGTTGGNAGCTCCGAGCCACACATGATG 1170
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 511 CATGATGACAAAGAGTGTGTCAGGCTTAAATGGTGTGAACAGGATTTCCACATGATG 570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1171 TCCCGGACCTTCATCCAGATCGACCGTGGCCAAACCCCTGGTGCAGCTTCCAGTGTGCGCATC 1230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 2302 AGCTGACGGCTTCCCGGCCCATCTGAGCGGCTGACCGTGAAGTCTCTCCGTGGG 2361
 Db 1708 AGAATTCGACGCTTTAGCCCTCTAAAGAGCCCTGACCATTCAGGTTCTTACTGTGGG 1767
 QY 2362 AGATGACACCCCGCGGTCGCTACTCTCT 2395
 Db 1768 AATGCCCTTCGACCTAAAATTAAATACACCTACT 1801

RESULT 11

US-10-097-580-2
 ; Sequence 2, Application US/10097580
 ; Publication No. US20030032168A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hirose, Kunitaka
 ; APPLICANT: Inoguchi, Eiji
 ; APPLICANT: Hakozaiki, Michinori
 ; APPLICANT: Ishioaka, Keiko
 ; APPLICANT: Matsushima, Kouji
 ; APPLICANT: Kuno, Kouji
 ; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
 ; FILE REFERENCE: Q57092
 ; CURRENT APPLICATION NUMBER: US/10/097,580
 ; CURRENT FILING DATE: 2002-03-15
 ; PRIOR APPLICATION NUMBER: 09/445,023
 ; PRIOR FILING DATE: 1999-12-03
 ; PRIOR APPLICATION NUMBER: JP 9-160422
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 2184
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: exon
 ; LOCATION: (1)..(2184)
 US-10-097-580-2

Query Match 22.0%; Score 627.2; DB 9; Length 2184;
 Best Local Similarity 60.8%; Pred. No. 4.8e-159;
 Matches 1079; Conservative 0; Mismatches 683; Indels 12; Gaps 3;

QY 631 AAGCGTTTCGTCTATCCCGCGGTACCTGGAGACGCTGGTGGCGGACGAGTCAATG 690
 Db 31 AAGCGATTGTGTCAGCGCCCGGTATGTGGAAACCATGCTTGTGGCAGACGAGTCGATG 90
 QY 691 GTCAAGTTCCACGCGCGGACCTGGAACATTATCTGCTGACGCTGGTGGTAAAGCGGCG 750
 Db 91 GCAGAAATTCACGCGCAGTGGTCTAAAGCAATACCTTCTCAGTTGTTTTCGGTGGCAGCC 150
 QY 751 CGACTCTACCGCCATCCGAGCATCCTCAACCCCATCAACATCGTGTGGTCAAGGTCGTG 810
 Db 151 AGATTGTACAACACCCCGACGATTCGTAATTCAGTTAGCCCTGGTGGTCAAGATCTTG 210
 QY 811 CTCTTAGAGATCTGACTCCGCGCCCAAGTCAACCGCATACCGGCTGACGCGCTCGC 870
 Db 211 GTCATCCAGTGAACAGAGCGGCGGAGTGACCTCCCAATGCTGCCCTCACTCTCGG 270
 QY 871 AACTCTGTGCTGGCAGAGAAGCTGAACAAAGTGAGTGACAAAGACACCCCGAGTACTGG 930
 Db 271 AACTTTTGCAACTGGCAGAGCAGCAGACACCCACCCAGTACCGGGATGAGAGCACTAT 330
 QY 931 GACACTGCCATCTCTTCAACGAGCAGGACCTGTGTGGAGCCACCACTGTGACACCCCTG 990
 Db 331 GACACGCAATCTTTTCAACGAGCAGGACTTGTGTGGTCCGAGACATGTGATCTCTT 390
 QY 991 GGCATGCTGATGGGTACCATGTGTGACCCCAAGAGAGCTGCTCTCATTTGAGGAC 1050
 Db 391 GGGATGCTGATGTGGAACTGTGTGATCCGAGCAGAGCTGCTCCGTATAGAAGAT 450

QY 1051 GATGGGCTTCCATCAGACCTTCAACACTGCCACAGCTGCCACAGCTGGCCACAGTGTTCACATGCCC 1110.
 Db 451 GATGGTTTTACAAGCTGCTTTCACACACGCCCATGAATTAGGCCACAGTGTGTTCACATGCCC 510
 QY 1111 CATGACAATGTGAAGTCTGTGAGGAGTGTGTTGGGAAGCTCCGAGCACAACACATGATG 1170
 Db 511 CATGATGATGCAAGCAGTGTGCCAGCTTAATGGTGTGAACGAGGATTTCCACATGATG 570
 QY 1171 TCCCGGACCTCATCCAGATCGACCCGTGCCAACCCCTGTGTCAGCTGCTGAGTGTGCTGCAATC 1230
 Db 571 GCGTCAATGCTTTCACACCTGGACACAGCCAGCCCTTGGTCTCTTTCAGTGTGCTTACATG 630
 QY 1231 ATACCGACTTCTTGGACAGCGGCGGACGCTGCTCTCTGAGCAACACCCAGCAAGCCC 1290
 Db 631 ATTACATCATTTCTGGATAATGTCATGGGAATGTTGATGGCAAGCTTCAGATCC 690
 QY 1291 ATCTCCCTGCCGAGGATCTGCCGGCGCCAGCTACACCTGAGCAGCAGTCCGAGCTG 1350
 Db 691 ATACAGCTCCAGGCGATCTCCCTGGCACCTTGTAGATGCCAACCGGACGCTGCCAGTTT 750
 QY 1351 GCTTTTGGCTGGGCTCCAAAGCCCTGTCTTACATGACGTAC--TGCACCAAGCTGTGG 1407
 Db 751 ACATTTGGGAGGACTCCAAACACTGCCCGATGTCAGCCAGCACATGTAGCACCTTTGTGG 810
 QY 1408 TGCACCGGGAAGCCCAAGGACAGATGGTGTGCCACAGCCGCCACTTCCCTGGGCGCAT 1467
 Db 811 TGTACCGGACCTCTGCTGGGTGGTGTGTCAACCAACAAACACTTCCCGTGGGCGCAT 870
 QY 1468 GGCACACGCTGTGGCGAGGCAAGCTCTGCTCAAAAGGCGCTGCGTGAGGAGACACAAC 1527
 Db 871 GGCACACGCTGTGGAGAGGGAATGGTGTATCAAGCGCAAGTGTGTGACAAACACCGAC 930
 QY 1528 CTCAACAGCACAGG-----GTGGATGTTCTCTGGGCAAAATGGATCCCTATGCCCC 1581
 Db 931 AGGAAGCATTTTCATACGCTTTTTCATGGAAGCTGGGACCATGCGGACCGTGGGAGAC 990
 QY 1582 TGCTCGGCACATGTGTGGGCGCTGACGCTGCCAGGAGGAGGAGTGCACCAACCCACC 1641
 Db 991 TGTTCGAGAACGTGGGTGGAGAGTCCAGTACACGATGAGGGAATGTGACACCCAGTC 1050
 QY 1642 CTGCGCAAGCGGGCAAGTACTGCGAGGAGTGAAGTGAATACCGATCTCTGCAATCTG 1701
 Db 1051 CCAAGATGAGGAGTACTGTGAAGCAACAGTGGCTACAGATCTCTGTAACTT 1110
 QY 1702 GAGCCCTGCCAGCTAGCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGAGGCTTC 1761
 Db 1111 GAGGACTGTCCAGAC---AATAATGGAACCCCTTTAGAGAGGAACAATGTGAAGACAC 1167
 QY 1762 AACGGCTCAACACACAGCACCACCGCTCACTCTGCCGCTGGCATGGTGGCCCAAGTAC 1821
 Db 1168 ACGAGTTTTCAAAGCTTCCCTTTGGGAGTGGGCTGGGATGGAATTCCTCAAGTAC 1227
 QY 1822 TCCGGCTGTCTCCCGGCAAGTCAAGCTCATCTGCCGAGCCAAATGGCACTGGCTAC 1881
 Db 1228 GCTGGCTCTCAACAAAGGACAGTGCAGCTCATCTGCCAAGCCAAAGCAATGGCTAC 1287
 QY 1882 TTCATGTGCTGCGCACCCCAAGTGTGTGGAGGAGCAGCTGTGCTCTCTGACTTCCACTCC 1941
 Db 1288 TTCCTGCTTTTCAGCCCAAGGTTGTTGATGTTACTTCCATGTAGCCCAAGTATCCACTTCT 1347
 QY 1942 GTCGTGTCCAAGGCAAGTGCATCAAGCTGGCTGTGATGGGAACCTGGGCTGCCAAGTAG 2001
 Db 1348 GTCGTGTGCAAGGACAGTGTGAAGCTGGTGTGATGCGCATCATAGACTTCCAAAG 1407
 QY 2002 AGATTCGACAAGTGTGGGTGTGTGGGGAGACAATAAGAGCTGCAAGAGTGAATGGA 2061
 Db 1408 AAGTTTGAATAATGTGTGTTTGGGGGAATGGATCTACTTGTAAAAAATATCAGCA 1467
 QY 2062 CTCCTTCAACAGCCCATGATGGCTACAAATTCGTGGTGGGCTGCCCCAGCGGCTCA 2121
 Db 1468 TCAGTTACTAGTCAACAAACCTGGATATCATGTCGCACAAATCCCAACTGGAGCCACC 1527


```
QY 1378 CCTATACAGTAC---TGCACCAAGCTGTGTGTCACCGGGAAGCCCAAGGACAGATG 1434
  || || || || || || || || || || || || || || || || || || || || ||
Db 1957 CCCGATCCAGCAGCAGATGTACACCTTGTGTGTACCGGCACCTCTGTGGGTGCTG 2016
  || || || || || || || || || || || || || || || || || || || || ||
QY 1435 GTGTGCAGACCCGACATTCCTCCCTGGCGGATGACACAGCTGTGGCGAGGCAAGCTC 1494
  || || || || || || || || || || || || || || || || || || || || ||
Db 2017 GTGTGTCAAAACAACTTCCCGTGGCGGATGACACAGCTGTGGAGAGGAAATGG 2076
  || || || || || || || || || || || || || || || || || || || || ||
QY 1495 TGCTCAAAAGGCGCTGCGTGGAGAGACAACTCAACAAG---CACAGGTGGAT 1548
  || || || || || || || || || || || || || || || || || || || || ||
Db 2077 TGTATCAACCGGAAGTGTGTGAACAAACCGACAGAAAGCATTTTATACGCTTTTCAT 2136
  || || || || || || || || || || || || || || || || || || || || ||
QY 1549 GGTTCCTGGCCAAATGGATCCTATGGCCCTCTGCGGCACATGTGTGGGGCGCTG 1608
  || || || || || || || || || || || || || || || || || || || || ||
Db 2137 GGAAGCTGGGAATGTGGGGCGCTGGGGAGACTGTTCAGAAACGTCGCGTGGAGGATC 2196
  || || || || || || || || || || || || || || || || || || || || ||
QY 1609 CAGTGGCCAGGAGGAGTGCACCAACCCCTGCCCCAACCGGGGCAAGTACTCGGAG 1668
  || || || || || || || || || || || || || || || || || || || || ||
Db 2197 CAGTACACGATGAGGGAATGTGAACCCAGTCCCAAGAAATGGAGGGAAGTACTGTGA 2256
  || || || || || || || || || || || || || || || || || || || || ||
QY 1669 GGAGTGAGGTGAAATACCGATCTGCAATCTGGAGCCCTGCCACAGCTCAGCCTCCGGA 1728
  || || || || || || || || || || || || || || || || || || || || ||
Db 2257 GGCAAACGAGTGGCTACACATCTGTAACTTGAAGACTGTCCAGAC---AATAATGA 2313
  || || || || || || || || || || || || || || || || || || || || ||
QY 1729 AAGAGCTTCGGGAGGAGCAGTGTGAGGCTTTCAAGGGTGTACCAACACAGCACCAACCGG 1788
  || || || || || || || || || || || || || || || || || || || || ||
Db 2314 AAAACCTTTAGAGAGGAACAATGTGAAGCACACAAAGAGTTTCAAAAGCTTCTCTTGG 2373
  || || || || || || || || || || || || || || || || || || || || ||
QY 1789 CTCACCTCCGCGGATGGTGCACCAAGTACTCCGGCTGTCTCCCGGACAAAGTGC 1848
  || || || || || || || || || || || || || || || || || || || || ||
Db 2374 AGTGGGCTCGGGTGAATGGATTCACAGTACGCTGGCTCTCACCACAAAGGACAGTGC 2433
  || || || || || || || || || || || || || || || || || || || || ||
QY 1849 AAGCTCATCTGCCGAGCAATGGCAGTGTCTTCTATGTGTCGACCAACAGGTGGTG 1908
  || || || || || || || || || || || || || || || || || || || || ||
Db 2434 AAGCTCATCTGCCAAGCAAGGATTTGGCTACTTCTCTTTTGCAGCCCAAGTTGTA 2493
  || || || || || || || || || || || || || || || || || || || || ||
QY 1909 GACGGCAGCTGTGCTCTCCTGACTCACCTCCGCTGTGTGTCCAAAGCAAGTGCATCAAG 1968
  || || || || || || || || || || || || || || || || || || || || ||
Db 2494 GATGGTACTCCATGTAGCCAGATTCACCTCTGTGTGTGCAAGGACAGTGTGTA 2553
  || || || || || || || || || || || || || || || || || || || || ||
QY 1969 GCTGGCTGTGATGGAACTGGGCTCAAGAAGAGATTCACAAAGTGTGGGTGTGGG 2028
  || || || || || || || || || || || || || || || || || || || || ||
Db 2554 GCTGGTGTGATGCATCATAGACTCCAAAGAGAGTTTGTATTAATGTGTGTGTTCCGGG 2613
  || || || || || || || || || || || || || || || || || || || || ||
QY 2029 GGAGACAATRAGACTCAAGAAGGTGACTGGACTCTTCAACCAAGCCCATGCATGCTAC 2088
  || || || || || || || || || || || || || || || || || || || || ||
Db 2614 GGAATGGATCTACTTGTAAAAAATATCAGGATCAGTTACTAGTGCACAAACCTGGATAT 2673
  || || || || || || || || || || || || || || || || || || || || ||
QY 2089 AATTTGCTGGTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCCGCAGCGGTTAC 2148
  || || || || || || || || || || || || || || || || || || || || ||
Db 2674 CATGATATCATCAAAATTCACACTGGAGCCACCAACATCGAAGTGAACAGCGGAACCCAG 2733
  || || || || || || || || || || || || || || || || || || || || ||
QY 2149 AAAGGCTGTATCGGGATGACAACTACTCTGCTCTGAAGAACAGCCCAAGCAAGTACCTG 2208
  || || || || || || || || || || || || || || || || || || || || ||
Db 2734 AGGGATCCAGGAACATAGCAGCTTTCTTGGCCATCAAGCTGCTGATGTCACATATAT 2793
  || || || || || || || || || || || || || || || || || || || || ||
QY 2209 CTCACAGGCAATTTGTTGGTGTGCGCGGTGAGCGGACCTGTTGTGTGAAGGCACTGTG 2268
  || || || || || || || || || || || || || || || || || || || || ||
Db 2794 CTTAATGGTGTACTACTTTGTCCACCTTAGAGCAAGACATTAATGACAAAGGTGTGTC 2853
  || || || || || || || || || || || || || || || || || || || || ||
QY 2269 CTGGGTGTACAGCGGACGGGACAGCGGTGGAGAGCCTGCAAGCTTCCCGGCCCATCTCTG 2328
  || || || || || || || || || || || || || || || || || || || || ||
Db 2854 TTGAGGTACAGCGGCTCCTCTGCGGCATTTGGAAGAAATTCGCAGCTTTAGCCCTCTCAA 2913
  || || || || || || || || || || || || || || || || || || || || ||
QY 2329 GAGCGCTGACCGTGGAGGTCTCTCGCTGGGGAAGATGACACGCGCCCGGGTCCGCTAC 2388
  || || || || || || || || || || || || || || || || || || || || ||
Db 2914 GAGCCCTTGACATCCAGGTTCTTACTGTGGGCAATGCCCTTCAGACCTAAATTAATAC 2973
  || || || || || || || || || || || || || || || || || || || || ||
QY 2389 TCCTTCT 2395
  || || || || || || || || || || || || || || || || || || || || ||
Db 2974 ACCTACT 2980
  || || || || || || || || || || || || || || || || || || || || ||
```

```
RESULT 14
US-09-918-171A-8
; Sequence 8, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 3638
; TYPE: DNA
; ORGANISM: Mus musculus ADAMTS-8
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (278)..(2992)
; NAME/KEY: misc_feature
; LOCATION: (3636)
; OTHER INFORMATION: n = T
US-09-918-171A-8

Query Match      20.7%; Score 590.4; DB 10; Length 3638;
Best Local Similarity 57.3%; Pred. No. 4.2e-149;
Matches 1334; Conservative 0; Mismatches 906; Indels 90; Gaps 11;

QY 157 CTCATTTTTCAGATCACAGCATTTTCAGAGGACTTTTACCTACACCTGACGCCGGATGCT 216
  || || || || || || || || || || || || || || || || || || || || ||
Db 440 CTCGCTTCCACCTCTCCGCTTTCGCCAGGCTTCGTGCTGCGCTGCGCTGACGCC 499
  || || || || || || || || || || || || || || || || || || || || ||
QY 217 CAGTTCCTGGCTCCGCTTCTCCACTGAGCATCTGGCGTCCCTCCAGGGCTCACC 276
  || || || || || || || || || || || || || || || || || || || || ||
Db 500 AGCTTCCTGGCGCGGAATTCAGATCAGGCGCTCGGGGCTCGAGCGCGCG 556
  || || || || || || || || || || || || || || || || || || || || ||
QY 277 GGGGCTCTTCAGACCTCGACGCTGCTTCTATTCTGGGGAGCTGAACGCCAGCGGAC 336
  || || || || || || || || || || || || || || || || || || || || ||
Db 557 GGGGCGAGCGGAGTGGTGGCTGCTTCTCTGSCACAGTGAATGAGAGACGGGAG 616
  || || || || || || || || || || || || || || || || || || || || ||
QY 337 TCGTTCGTGTGTGAGCCTGTGCGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCC 396
  || || || || || || || || || || || || || || || || || || || || ||
Db 617 TCGCTGGCGCGATGAGTGTGTGCGGGCTGGAGCGCTCGTCTTCTGCGCAGCGGAG 676
  || || || || || || || || || || || || || || || || || || || || ||
QY 397 GAGTATGTCATTAGCCCGCTGCCAATGCTAGCGCGCGCGGCGCAGCAACAGCCAG 456
  || || || || || || || || || || || || || || || || || || || || ||
Db 677 GAGTTCACCATCCAGCCAGGCGCTGGGACTCCCTGGACCGCTCATGCGCTCGAG 736
  || || || || || || || || || || || || || || || || || || || || ||
QY 457 GCGGCACACCTTCTCAGCGCGCGGGTGTTCGGCGGGGCTTCGGGAGA----CCCCAC 512
  || || || || || || || || || || || || || || || || || || || || ||
Db 737 CGCTGGGGCGGGAGACGCGCGCGAAGACCCCGGCTCGCTGCGCGCAAGTTTTCGCC 796
  || || || || || || || || || || || || || || || || || || || || ||
QY 513 CTCTCGCTGCGGGTGGCTCGGGCTGGAAACCCGCCATTCCTACGGGCCCTGGACCTTA 572
  || || || || || || || || || || || || || || || || || || || || ||
Db 797 CTCCCTCAAGGACTGGAGTGGAGGTGGATGGTAATGGGCGAGGACAGAGAGAAAGT 856
  || || || || || || || || || || || || || || || || || || || || ||
QY 573 CAAGCGCGGGCGGGCTTCGGGGAGAGTGTAGCGCGCGCGAGCTCTGGG-----624
  || || || || || || || || || || || || || || || || || || || || ||
Db 857 GACAAGAGAGGACAAGAACAGCAGGAGGGGTTGCTCAAAGAGACAGAGACTCC 916
  || || || || || || || || || || || || || || || || || || || || ||
QY 625 -----CGCGCAAGCGTTTCGTGTCTATCCCG 651
  || || || || || || || || || || || || || || || || || || || || ||
Db 917 CGCAAGTCCACACCTTCGATCCAAACTAGAACAGAGAGTGTGTGTCAGGCT 976
  || || || || || || || || || || || || || || || || || || || || ||
QY 652 CGGTACGTGGAGCGCTGGTGTGTCGGGACGAGTCAATGGTTCACGCGCGGAC 711
  || || || || || || || || || || || || || || || || || || || || ||
Db 977 CGCTTCGTGGAACACATTCGTGTGCTGATGCTCATGTGCTTATGATGGACCGAC 1036
  || || || || || || || || || || || || || || || || || || || || ||
QY 712 CTGGAACATTTATCTGTGACGCTGTGGCAACGCGCGCGGCTCTACCGCATCCAGC 771
  || || || || || || || || || || || || || || || || || || || || ||
```

[illegible]

Db	2111	AATGGCTACAACCA--CACTGACCTGGATGGGAATTTCTCTCAGTGGGTCCCCAAGTAT	2167
QY	1822	TCCGGCTGTCTCCCCGGGACAGATGCAAGCTCATCTGCCAGACCAATGGCACTCGCTAC	1881
Db	2168	TCAGAGTGTCTCCCCCGAGACGATGATCAAGCTGTTTTTGAGAGCCCGTGGGAGGACGTAG	2227
QY	1882	TTCTATGTGCTGGCACCAAGGTGGTGGACGGCAGCTGTGCTCTCTGACTCCACCTCC	1941
Db	2228	TTCAAGTGTTTGAAGCTAAGGTGATCGATGGCACTCTGTGTGGACCGGATACTGTGTC	2287
QY	1942	GTCGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGATGGGAACCTGGGCTCCACGAAG	2001
Db	2288	ATCTGCGCTGGGGGCAATCTGTTAAGGCTGGCTGTGACCATGTGTGAAGTCACTAAAG	2347
QY	2002	AGATTGACAAAGTGTGGGTGTCTGGGGGAGACAATAAGAGAGTGCACGAAGGTGATGGA	2061
Db	2348	AAGCTGGACAAATGTGGGCTGTGTGGGGCAAGGCACCTGGCTGTAGGAAGATCTTCGGT	2407
QY	2062	CTCTTACCAAGCCCATGCATGCTACAAATTTCTGTGTGGGCATCCCGCAGGCGCTCA	2121
Db	2408	TCCTTTCACCCCTTCAGTTATGGCTACAAATGACATTCAGTCACCATCCCAAGTGGTGGCACA	2467
QY	2122	AGCATGCACATCCGCCAGCGCGTGTACAAGGSGCTCATGGGGATGACAACATCTCGCT	2181
Db	2458	AACATTCATGTGAACAGCGGGATACCCAGGGGTGAGAACGCGGAGTAGCTCGGCG	2527
QY	2182	CTGAAGAACGCCAAGCAAGTACCTGCTCAACGGGCGATTTCTGGTTCGGCGGTGGAG	2241
Db	2528	CTGAAGACAGCAATGGCAGTACCTGCTCAATGTTAACTTGGCCATCTCTTGCCATAGAG	2587
QY	2242	CGGACCTGGTGTGAAGGCACTGTGCTGCGGTACAGGGCAGGGCACAAGCGGTGGAG	2301
Db	2588	CAAGACATCTGTGTGAAGGGGACATCTCTGAAGTACAGTGGCTCCATGGCTACCGCTGGAG	2647
QY	2302	AGCCTGCAGGCTTCCCGGCCATCTCTGGAGCGCTGACCGTGGAGGTCTCTCTCGCTG--	2358
Db	2648	CGGCTGCAGAGCTTCCAGGCCCTGCCTGASCCCTTTACGTACAGCTCCTGACTCTGTCT	2707
QY	2359	GGGAAGATGACACGGCCCCGGGTCCGCTACTCCTTCTATCTGCCCAAGA	2408
Db	2708	GGTGAGTCTCTCCCTCAAAAGTCAGATATACCTCTCTTTGTGCCCAATGA	2757

RESULT 15
US-10-174-590-351

US-10-174-390-331
; Sequence 351, Application US/10174590
; Publication No. US20030008352A1

; PUBLICATION NO: 0320
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Baker, Kevin
APPLICANT: Chen, Jian

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers

APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey

;
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul

APPLICANT: Godowski, Paul

APPLICANT: Gurney, Austin

APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin

; APPLICANT: Wood, William I

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRET:

; TITLE OF INVENTION: ACIDS

; FILE REFERENCE: P3430R1C42

; CURRENT APPLICATION NUMBER

; CURRENT FILING DATE: 2002

; Prior application removed

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 351

; CSEQ ID NO 331
; LENGTH: 4407

LENGTH: 4407
; TYPE: DNA

TYPE: DNA
ORGANISM: Homo Sapien

ORGANISM: HOMO sapien
US-10-174-590-351

US-10-174-390-351

Query Match 20

Query Match 20.1

[illegible]

Search completed: May 16, 2003, 00:52:32Z
Job time : 382 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 19:42:09 ; Search time 107 Seconds
(without alignments)
8177.092 Million cell updates/sec

Title: US-09-965-631-3
Perfect score: 2853
Sequence: 1 atgtcttctgtggcctct.....gcgtctctgagccgtgctgta 2853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15333831 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	626.6	22.0	4676	4	US-09-130-491-1
2	617.2	21.6	4858	4	US-09-392-184-1
3	611.4	21.4	3706	4	US-09-484-970B-58
4	598.2	21.0	3126	4	US-09-392-184-7
5	590.4	20.7	3638	4	US-09-369-364A-8
6	583.2	20.4	4192	4	US-09-122-126B-1
7	549	19.2	3250	4	US-09-122-126B-14
8	547.4	19.2	3002	4	US-09-369-364A-1
9	469.8	16.5	2625	4	US-09-369-364A-14
10	409.6	14.4	5804	4	US-09-369-364A-12
11	318	11.1	2114	4	US-09-130-491-7
12	229.4	9.1	7739	4	US-09-369-364A-10
13	245.6	8.6	5357	4	US-09-392-184-5
14	238.6	8.4	3218	4	US-09-369-364A-6
15	217.4	7.6	1520	4	US-09-369-364A-3
16	212.8	7.5	3885	4	US-09-369-364A-16
17	185.4	6.5	2450	4	US-09-491-522-2
18	185.4	6.5	6692	4	US-09-491-522-1
19	177.8	6.2	2450	4	US-09-491-522-9
20	177.8	6.2	4580	4	US-09-491-522-8
21	162.2	5.7	703	4	US-09-392-184-6
22	116	4.1	3675	4	US-09-392-184-3
23	116	4.1	4042	4	US-09-392-184-2
24	91.6	3.2	2848	4	US-09-369-364A-4
25	77	2.8	2023	4	US-09-491-522-6
26	77	2.7	385	4	US-09-392-184-23
27	67.6	2.4	349	4	US-09-392-184-24

28	61.4	2.2	1642	4	US-09-369-364A-18	Sequence 18, Appl
29	57	2.0	1476	4	US-09-930-872-1	Sequence 1, Appl
30	55.8	2.0	1803	4	US-09-369-364A-20	Sequence 20, Appl
31	50.4	1.8	657	3	US-08-985-526-2	Sequence 2, Appl
32	50.4	1.8	1326	3	US-08-985-526-4	Sequence 4, Appl
33	49.4	1.7	716	4	US-09-276-531-117	Sequence 117, App
34	47.4	1.7	3014	2	US-08-808-982-1	Sequence 1, Appl
35	47.4	1.7	3014	4	US-09-306-902A-1	Sequence 1, Appl
36	47.2	1.7	1173	3	US-08-706-216-5	Sequence 5, Appl
37	46.4	1.6	1101	4	US-09-152-060-23	Sequence 23, Appl
38	44.6	1.6	502	4	US-09-392-184-12	Sequence 12, Appl
39	44	1.5	7452	3	US-08-592-500-1	Sequence 1, Appl
40	44	1.5	7452	3	US-08-195-006-1	Sequence 1, Appl
41	44	1.5	7452	5	PCT-US94-07644A-1	Sequence 1, Appl
42	43.6	1.5	2063	1	US-08-715-554-1	Sequence 1, Appl
43	43.6	1.5	2063	1	US-08-583-118-1	Sequence 1, Appl
44	43.4	1.5	7218	2	US-08-232-463-14	Sequence 14, Appl
45	43	1.5	20235	1	US-07-642-734C-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-130-491-1
; Sequence 1, Application US/09130491
; Patent No. 6416974

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

; FILE REFERENCE: 09404/041001

; CURRENT APPLICATION NUMBER: US/09/130,491

; CURRENT FILING DATE: 1998-08-07

; EARLIER APPLICATION NUMBER: US 60/058,108

; EARLIER FILING DATE: 1997-09-05

; EARLIER APPLICATION NUMBER: US 60/054,961

; EARLIER FILING DATE: 1997-08-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 4676

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (460)....(3360)

US-09-130-491-1

Query Match 22.0%; Score 626.6; DB 4; Length 4676;

Best Local Similarity 57.0%; Pred. No. 1.1e-131;

Matches 1304; Conservative 0; Mismatches 914; Indels 69; Gaps 6;

QY 175 GCATTTTCAGAGACATTTTACCTACACCTGACCGCGGAGTCTCAGTTCCTGGCTCCGCG 234

Db 697 GCCTTTTACACAGCAGCTGGATCTGGAGCTGCGCGCCGACAGCAGCTTTTGGCGCGCG 756

QY 235 TTCTCCACTGAGCATCTGGCGTCCCTCCAGGGCTCACCAGGGGCTC---TTCAGAC 291

Db 757 TTCACGCTCCAGAACGTTGGGGCGCAATCGGGTCCGAGACGCCGCTTCCGGAACACGAC 816

QY 292 CTGCGACGCTGCTTCTATTCTGGGACGTGAACCGCGCGGACCTGCTTGGCTGTGTG 351

Db 817 CTGGCGCACTGCTTCTTACTCCGSCACCGTGAATGGGATCCAGCTCGGCTCGCGGCTC 876

QY 352 AGCTGTGGGGGGGCTCCGCGAGCCTTTGGCTACCGAGGCGCGAGTATGTCATTAGC 411

Db 877 AGCTGTGGGGGGGCTGCGCGCGCTTCTTACTGCTGGGGAGCGGTATTCATCCAG 936

QY 412 CGCTCCCAATGCTAGCGCGC---CGGCGCGGCGGACGACACACCCAGGGGCGCACCTTCT 470

Db 937 CGCTGCGCGCGGCGGACGCGCTCGCCACCGCGCGGCGGAGGAGGAGGAGGAGGAGGAG 996

QY 471 CCAGCCGCGGCTTCCGGGGGCGCTTCCGAGACCCCACTCTCGTGGGGGTGC 530
Db 997 CCACACAGTTCCACCTCTCGGGGGAATGCGAGGCGAGTAGGCGGCACGTGCGGG 1056
QY 531 CTCGGGCTGGAACCCCGCATCTAGGGGCGCTGAGCCCTTACAGCCCGCGGGGGGG 590
Db 1057 GTCGTGAGACAGAGCCCGGGCGACTGGGAAAGGAGAGAGAGAGAGAGAGAG 1116
QY 591 CTTTCGGGAGAGTCGTAGCCGGCGCAGGTCGGGCGC----- 627
Db 1117 ACTGAGGGCGAGACGAAGGGCTCAGTGTGCGCCGAGAGACCCGGCACTGCAAGCGGTA 1176
QY 628 -----GCCAAGCGTTTCTGTCTATCCTGCGGAGTAC 657
Db 1177 GGACAGCCACAGAACTGGAAGCATAAGAAAGAACGAGATTTGTGCCAGTCAACCGCTAT 1236
QY 658 GTGAGAGCGTGTGTGCGGGAGAGTCAATGGTCAATTTCCAGGGCGGACCTGGNA 717
Db 1237 GTGGAACCATGCTTGTGGCAGACCACTGATGGCAGAAATCCACGGCAGTGGTCTAAAG 1296
QY 718 CATTATCTGTGACGTGCTGGCAACGCGGCGGACTCTACCGCATCCACAGCATCCPC 777
Db 1297 CATTACCTCTACGTGTTTTCGGTGGCAGCAGATTTGTAACACACCCAGCAATCTGT 1356
QY 778 AACCCCATCAACATCTGTGTGTCAGAGTCTCTCTTCTTAGAGATGCTGACTCCGGGCC 837
Db 1357 AATTCACTTACGCTGTGTGTGTAAGATCTTGGTCTATCCAGTCAACAGAAAGGGCGG 1416
QY 838 AAGTCAACCGCAATGCGGCCCTGAGCTGCGCAACTTCTGTCCCTGGGAGAGAGCTG 897
Db 1417 GAATGACCTCCAAATGCTGCCCTCACTCTCGGAATTTTGAACACTGGCAGAGAGCAC 1476
QY 898 AACAAGTGAAGTGAACAGACCCCGAGTACTGGGACACTGCCATCTCTTCCACAGGCAG 957
Db 1477 AACCCACCACTGACCGGATGACAGCACTATGACACAGCAATCTTTTCACAGACAG 1536
QY 958 GACTGTGTGGAGCCACCACTGTGACACCCCTGGGATGGCTGATGGGTACCATGTGT 1017
Db 1537 GACTGTGTGGTCCAGACATGTGATCTCTTGGGATGGCTGATGTTGGAATGTGTGT 1596
QY 1018 GACCCAGAGAGCTGCTCTGTCATTGAGGAGGATGGGCTTCATCAGCTTCACCACT 1077
Db 1597 GATCCGAGAGAGCTGCTCCGTCATAGAGATGATGTTTACAGCTCCCTTCACACA 1656
QY 1078 GCCCAGAGCTGGGCGACGCTTCAACATGCCCATGCAATGAAATCTGTGAGGAG 1137
Db 1657 GCCCATGAATTAGGCCAGGTTTAACTGCCACATGATGATCAAAAGCAGTGTGCCAGC 1716
QY 1138 GTGTTGGGAGCTCGAGCCACACAGATGTCCCGACCTCATCCAGATCGACCTG 1197
Db 1717 CTTAATGGTGAACAGGATTCACATGATGGCGTCAATGCTTTCCAACTGGACCA 1776
QY 1198 GCCAACCCCTGTGACGCTGCTGCTGATCATCAGGATCTCTGACAGCGGGCAC 1257
Db 1777 AGCAGCTTGTGCTCTTCCAGTGGCTAGATGATGATGATGATGATGATGATGAT 1836
QY 1258 GGTGACTGCTCTGAGACCAACCAAGCCATCTCCCTGCGGAGGATGTCGGGGG 1317
Db 1837 GGGGAATGTTGATGGACAAAGCTTCAGATCCCATACAGTCCCAAGCTCTCCCTGGC 1896
QY 1318 GCCAGCTACACCTGAGCAGAGTGGAGCTGCTTTGGCGTGGGCTCCAAAGCCCTGT 1377
Db 1897 ACCTCGTACAGTGCACCGGAGTGGAGTTCATTTTGGGAGGACTCCAAACACTG 1956
QY 1378 CTTTACATGAGTAC---TGACCAAGCTGTGTGTGACCGGAGGAGGAGGAGGAGATG 1434
Db 1957 CCCGATGACCCAGCAGCATGACACCTTGTGTGTACCGGACCTCTGTGTGGGTGCTG 2016
QY 1435 GTGTGCGAGACCCCGCACTTCCCTGGGCGATGGCACACAGCTGTGGCAGGCGCAAGCTC 1494
Db 2017 GTGTGTCAAAACCAACTTCCCTGGGCGGATGGCACCAGCTGTGGAGAGGAAATG 2076
QY 1495 TGCCTCAAGGGGCGCTCGTGGAGAGACACCACTCAACAAG-----CACAGGTGGAT 1548

Db 2077 TGTATCAACGCAAGTGTGTGAACAAAACCCACAGAAGCAATTTTGATACGCTTTTCAT 2136
QY 1549 GGTTCCTGGGCAAAATGGGATCCCTATGTCCTGTCGCGCACATGTGTGGGGCGTG 1608
Db 2137 GGAAGCTGGGAATGTGGGGCTTGGGAGACTGTTTCGAGAAGCTGCGGTGGAGGAGTC 2196
QY 1609 CAGTGGCCAGGAGGAGTGCACCAACCCACCCCTGCCAACCGGGGCAAGTACTCGGAG 1668
Db 2197 CAGTACAGATGAGGAATGTGACAAACCCAGTCCCAAGAATGGAGGAAAGTACTGTGAA 2256
QY 1669 GAGTGAAGGTGAATACCGATCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCGGA 1728
Db 2257 GGCAAACAGTGCCTACAGATCTGTAACTTGGAGACTGTCCAGAC---AATAATGGA 2313
QY 1729 AAGAGCTTCCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTACAAACACAGCACCACCGG 1788
Db 2314 AAAACCTTTAGAGAGAACATGTGAAGCACACAAACAGTTCCTTTGGG 2373
QY 1789 CTCATCTCGCGGTGGCATGGGTGCCCAAGTACTCCCGCGTGTCTCCCGGGACAAAGTGC 1848
Db 2374 AGTGGCGCTGGTGAATGATTCCCAAGTACGCTGGCGTCTCACCAGAGCAGGTGC 2433
QY 1849 AAGCTATCTCCGAGCAATGGCACTGGTACTTCTATGTGTGGCACCACCAAGTGTG 1908
Db 2434 AAGCTATCTCCGAGCAATGGCACTGGTACTTCTTTCGTTTGGAGCCCAAGTGTGTA 2493
QY 1909 GACGCGCTGTCTCTCTGCTGCTCCAGCTCCCTGCTGTGTGCCAAGGCAAGTGCATCAAG 1968
Db 2494 GATGGTACTCCATGTAGCCAGATTCACCTCTGCTGTGTGCAAGGACAGTGTGTA 2553
QY 1969 GCTGCTGTGATGGAGCACTGGGTGCCAAGAGATTCGCAAGTGTGGGGTGTGGG 2028
Db 2554 GCTGCTGTGATGCGATCATAGACTCCAAAAGAGTGTGATAATGTGTGTGTGGGG 2613
QY 2029 GGAGACATAGAGTGCAGAGGTGACTGACTCTTCAACAGCCCATGCTATGCTAC 2088
Db 2614 GGAATGATGATCTTGTGAAAAAATATCAGGATCAGTTACTGTGAAAAACCTGTGATAT 2673
QY 2089 AATTTCGTGGTGGCATCCCGCAGGCGCTTCAAGCATCGACATCCCGCAGCGGGTTAC 2148
Db 2674 CATGATATCATCAATTTCCAACTGGAGCCACCAACATCGAAGTGAACAGCGGAACAG 2733
QY 2149 AAGGGCTGTGCGGATGACAACTACTGCTGTGAAGACACCAAGCAAGCACTG 2208
Db 2734 AGGGATCCAGGAACAATGGCAGCTTTCTGCCATCAAAGCTGCTGATGGCACAATATT 2793
QY 2209 CTCAGGGGCAATTCGTGTGCTGCGCGGTGAGCGGACCTGGTGTGAAGGCGAGTCTG 2268
Db 2794 CTTAATGTGACTACACTTTGTCACCTTAGAGCAAGACATATGTACAAAGTGTGTGTC 2853
QY 2269 CTGCGGTACAGCGGACCGGACAGCGGTGAGAGCCTGAGAGCTTCCCGGCCCATCTG 2328
Db 2854 TTGAGTACAGCGCTCTCTCGCGCATTTGGAAGAATTCGACGCTTTAGCCCTCTCAA 2913
QY 2329 GAGCGCTGAGCGGTGAGGCTCTCTCGTGGGAGATGACACCCCGGCTCCGCTAC 2388
Db 2914 GAGCGCTGAGCATCCAGGTTCTTACTGTGGCAATGCCCTTCGACCTAAATTAATAC 2973
QY 2389 TCCTTCT 2395
Db 2974 ACCTACT 2980

RESULT 2

US-09-392-184-1
; Sequence 1, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PROTEASE HOMOLOGS
; FILE REFERENCE: 5800-55

```

; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4858)
; OTHER INFORMATION: reprotolysin (ADAM family of metalloprotease)
; NAME/KEY: misc_feature
; LOCATION: (1)...(4858)
; OTHER INFORMATION: n = A,T,C or G
US-09-392-184-1

Query Match          21.6%; Score 617.2; D3 4; Length 4858;
Best Local Similarity 57.0%; Pred. No. 1.5e-129;
Matches 1305; Conservative 0; Mismatches 913; Indels 70; Gaps 7;

QY 175 GCATTTCAGGAGGACTTTTACCTACACCTGACGCGCGGATGCTCAGTTTGGCTCCCGCC 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 GCCTTTGACACGAGCTGGATCTGAGCTGGCGCGCCGACAGCAGCTTTTGGCGCCCGC 750

QY 235 TTCTCCACTGAGCATCTGGCGTCCCTCCAGGGGCTACCGGGGCTC---TTGACAG 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 TTCAGCTCCAGACGCTGGGCGCAATCCCGGTCGCGAGACGCCGCTTCGGAACCGAC 810

QY 292 CTGCGACGCTGCTTATTTCTGGGACGCTGAACCGCGAGCGGAGCTGCTTGGCTGCTGTG 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 CTGGCGCACTGCTTCTACTCCGGCACCGTGAATGGCGATCCACGCTGGGTGGCGCCCTC 870

QY 352 AGCTGTGCGGGGGCTCCCGGAGCCTTTGGCTACCGAGCGCGGATGTCATATAGC 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 871 AGCTCTCGGAGGCGTCCGGGGGCTTCTACCTGCTGGGGAGGCGTATTTTCATCAG 930

QY 412 CGCTGCCCAATGCTAGGCGC-CGGGGCGCAGCGCAACAGCCAGCGGCGCACACTTCT 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 931 CGCTGCGCGCGCCAGCGAGCGCTCGCCACCGCGCGCCGAGGAGAAAGCGCGCGCA 990

QY 471 CCAGCGCGGGGTGTTCGGGCGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGC 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 991 CCATACAGTTCACCTCTCTCGCGGAAATCGCGAGGCGACGTAGCGGCGACGTGCGGG 1050

QY 531 CTCGGGTGGAACCCCGCATCTACGGGCGCTGGACCTTTACAGCGCGGGGGCGGG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1051 GTCGTGGACGACGAGCGCGCGCTGGGAAAGCGGAGACCGAAGACGAGGCGAAGG 1110

QY 591 CTTGGGGGAGAGTCTGAGCGGCGCAGTCTGGGCGC----- 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1111 ACTGAGGCGAGGAGGCGCTCAGTGTGCGCGCGAGGACCGGCGACCTGCAAGGCGTA 1170

QY 628 -----GCCAAGCGTTTGGTGTCTATCCCGCGGTAC 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1171 GGACAGCCACAGGAAGCTGGAAGCATAGAAGAGCGATTTGTGTCATCCACCGGTAT 1230

QY 658 GTGAGAGCGTGTGTGGGAGCAGTCAATGTCAAGTTCACAGCGCGGAGACCTGGA 1717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1231 GTGAAACCAATGCTGTGGGACGACGTCGATGGCAGAAATTCACAGCGCTGTCTAAAG 1290

QY 718 CATTATCTGCT-GAGCGTGTGGCAACGCGGCGGCGACTCTACCGCCATCCAGCATCT 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1291 CATTACCTTCTCAACGTTGTTTGGGTGAGGCGAGATTTACAAACAOCCAGCATTCG 1350

QY 777 CAACCCCATCAACATCGTTGTGTCGAAGTGTCTTCTTTAGAGATGCTGACTCCGCGGC 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1351 TAATTCAGTGTAGCTGTGGTGGTGAAGATCTTGGTCATCCAGATGAACAGAGGGGCC 1410

QY 837 CAAGGTCAACGGCAATCGGCGCTGACGCTGCGCAACTTCTGTGCTGGCAGAGAGCT 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1411 GGAAGTGACCTCCAACTGCTGCTCTCTGCGGAACTTTTGCAACTGGCAGAGCAGA 1470
```

```

QY 897 GAACAAAGTGAAGTGAAGCAAGCAAGCAAGTACTGGGACACTGCCATCTCTTCAACAGCA 956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1471 CAACCCACCCAGTGAACCGGATGACAGAGCACTATGACACAGCAATCTTTTCAACAGCA 1530

QY 957 GGACTGTGTGGAGCCACCACTGTGACACCCCTGGGATGCTGATGTGGGTACCATGTG 1016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1531 GGACTTGTGTGGGTCCAGACATGTGATCTCTTGGGATGCTGATGTGTGGAATGTGTG 1590

QY 1017 TGACCCCAAGAGAGTGTCTCTGTCATTGAGGACGATGGCTTCCATCAGCCTTCAACCAC 1076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1591 TGATCCGAGCAGAAGTGTCTCCGTCATAGAGATGATGTTTACAGCTGCTTCAACCAC 1050

QY 1077 TGCCACGAGCTGGGCGACGCTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAGGA 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1651 AGCCATGAATTAGGCCACGCTTTAAACATGCCACATGATGATGCAAGCAGTGTGCCAG 1710

QY 1137 GGTGTTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGACCCCTCATCCAGATCCAGCG 1196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1711 CTTAATGTGTGAACCAAGGATTCACACATGATGGGCTCAATGCTTTTCCAACTTGACCA 1770

QY 1197 TGCCAAACCCCTGTCAGCCTGTCAGTGTGCTCATCATCACCGACTTCTTGACAGCGGCA 1256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1771 CAGCCAGCCTTGTCTCTTTCAGTGCCTACATGATTAATCATATTTCTGGATTAATGCTCA 1830

QY 1257 CGGTGACTGCTCTCTGACCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGG 1316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1831 TGGGAAATGTTGATGGACAAGCTCAGAACTCCATACAGCTCCAGGCGATCTCCCTGG 1890

QY 1317 CGCCAGCTACACCCCTGAGCCAGCAGTGCAGCTGGCTTTTGGGCTGGGCTCCAGCCCTG 1376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1891 CACCTGTAGATGTCGAACCGGCGAGTGCAGTTTACATTTGGGGAGGACTTCCAAACACTG 1950

QY 1377 TCCTTATACATCAGTAC---TGCCACCAAGCTGTGTCACCGGAGGCGCAAGGACAGAT 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1951 CCCTGATGACGCCAGCAGACATGTAGCACTTGTGTGTACCGGCACTCTGTTGGGTGCT 2010

QY 1434 GGTGTGCGAGACCCCGCACTTCCCTTGGGCGGATGACACAGCTGTGGCGAGGCGAAGCT 1493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2011 GGTGTGTCAAAACAACTTCCGCTGGGCGGATGTCACAGCTGTGGAGAAGGAAATG 2070

QY 1494 CTGCTTCAAGGGCGCTGGGTGGAGAGACACACCTCAACAG-----CACAGGTGGA 1547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2071 GTGTATCAACGGCAAGTGTGTGAACAAACCGAGACAGATTTTGTATACGCTTTTCA 2130

QY 1548 TGTCTCTTGGGCAAAATGGGATCCCTTATGGCCCTCTCGCGCACATGTGGTGGGGCGCT 1607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2131 TGAAGCTGGGAAATGTGGGGCGCTTGGGAGACTGTTGAGAGACGTGCGGTGGAGAGT 2190

QY 1608 GCAGCTGGCCAGGAGCAGTGCACAAACCCCTGCGCAACGGGGGCAAGTACTGCGA 1667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2191 CCAGTACAGATGAGGGAATGTGACAAACAGTCCCAAGAAATGGAGGGAAGTACTGTGA 2250

QY 1668 GGGAGTGGGTGAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGCG 1727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2251 AGCAACAGGTGCGCTACAGATCTGTAACTTTAGGACTGTCCAGAC---AATTAATGG 2307

QY 1728 AAAGACTTTCCGGGAGGAGCAGTGTGAGGCTTTTCAACGCTTACAAACAGCAGCAACCG 1787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2308 AAAAAGCTTTAGAGAGAAACATGTGAACACACACAGAGTTCCTCAAAAGCTTCTTTGG 2367

QY 1788 GCTCACTTCGCGGTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGCAAGT 1847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2368 GAGTGGGCTGGGTGAATGGAATTCCTCAAGTACGCTGCGCTCACCAGGAGCAGGTG 2427

QY 1848 CAAGCTCATCTCCGAGCAATGGCACTGGCTACTTCTATGTGCTGGCACCCCAAGTGTGT 1907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2428 CAACTCATCTGCCAAGCCAAAGGCAATGGCTACTTCTTCTGTTTTCGAGCCCAAGTGTGT 2487

QY 1908 GGACGCGACGCTGTCTCTCTCTGACTTCCACTCCCGTGTGTCGCAAGCACTTCCATCA 1967
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2488 AGATGGTACTCCATGTAGCCCAAGTTCACCTCTGCTGTGTGCAAGGACAGTGTGTAA 2547

QY 1968 GGCTGGCTGTGATGGGAACCTGGGCTCCAGAAAGAGATTCGACAAAGTGTGGGTGTGTG 2027
```

Db 2548 AGCTGGTTGTGATCGCATATAGACTCCAAAGAAAGTTTGTATAAATGTGGTGTGGCG 2607
QY 2028 GGGAGACAATAAGAGCTGCAAGAGTGTGACTGGACTCTTCCACCAAGCCCATCATGGCTA 2087
Db 2608 GGGAAATGGATCTACTTGTAAAGAAATATACAGATCAGTTACTAGTCAAAAACCTGGATA 2667
QY 2088 CAATTTCTGTGGCATCCCGCAGCGCTCAAGCATCGACATCCGCTAGCGGGTTA 2147
Db 2668 TCATGATATCATACAAATCCAACTGGAGCCACCAATCGAAGTGAACACGGGNACCA 2727
QY 2148 CAAGGCGCTGATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGGCAAGTAAC 2207
Db 2728 GAGGGATCAGGAACAATGGCAGCTTCTTGCCATCAAGCTGCTGATGGACATATAT 2787
QY 2208 GCTCAAGCGGCATTTCTGTGTGTGCGGGTGGAGCGGGACCTGGTGGTGAAGGGCAGTCT 2267
Db 2788 TCTTAATGGTGACTACACTTTTGTCCACCTTAGAGCAAGACATATATACAAAGGTGTTGT 2847
QY 2268 GCTGCGGTACAGCGGCACGGGCACAGCGGTGGAGAGCCTGACGGCTTCCTCGGCCATCCT 2327
Db 2848 CTTGAGGTACAGCGCTCCTCTCGGCATTTGGAAAGNAATTCGACGCTTAGCCCTCTCAA 2907
QY 2328 GGAGCCGCTGACCGTGGAGTCTCTCCGHHGGGAAGATGACACCGCCCGGGTCCGCTA 2387
Db 2908 AGAGCCCTTGACCATCCAGGTTCTTACTGTGGCAATGCCCTTCGACCTAAAATTAATA 2967
QY 2388 CTCCTTCT 2395
Db 2968 CACCTACT 2975

RESULT 3

US-09-484-970B-58
; Sequence 58, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484, 970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 58
; LENGTH: 3706
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 007074.1
; NAME/KEY: unsure
; LOCATION: 3634, 3638-3639, 3642-3643, 3647-3648, 3652, 3654-3658, 3664, 3674, 3681
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-58

Query Match 21.4%; Score 611.4; DB 4; Length 3706;

Best Local Similarity 60.6%; Pred. No. 2.7e-128;

Matches 1076; Conservative 0; Mismatches 686; Indels 13; Gaps 4;

QY 631 AAGCGTTTCGTGTCTATCCCGCGGTAGCTGGAGACGCTGGTGGTTCGGGACGAGTCAATG 690
Db 183 AAGCGATTGTGTCAGTCAACCGCTATGTGGAACCATGCTTGTGGCAGACCACGATG 242
QY 691 GTCAGTTCACGCGCGGACCTGGACATTAATCTGTGACGCTGCTGGCAACGCGGCG 750
Db 243 GCAGAAATCCACGCGGTGGTCTAAAGCAATACCTTCTCAGTTGTTTCGGTGGCAGCC 302
QY 751 CGACTCTACCGCCCATCCAGCATCCTCAAGCCCATCAACATCGTTGTGGTCAAGGTGCTG 810
Db 303 AGATTGTACAAACACCCACGACATTCGTAAATTCAGTTAGCTAGCTGGTGGTGAAGATCTG 362

QY 811 CTTCTTAGAGATCGTACTCGGGCCCCAAGGTCCACGGCAATGCGGCCTGAGCCTCGCG 870
Db 363 GTCATCCACGATGAACAGAGGGCGGAGTGTACCTCCAAATGTGCGCTCACTCTCGGG 422
QY 871 AACTTCTGTGCTGGCAGAGAAAGCTGAACAAAGTGTGAGTACAGACACCCCGAGTACG 930
Db 423 AACTTTTGAATGGCAGAGCAGCAACACCCACCGGATGACGGGATGAGAGACATAT 482
QY 931 GACACTGCATCTCTTCAACAGCAGGACCTGTGTGGAGCCACACCTGTGACACTG 990
Db 483 GACACAGCAATCTCTTCCACAGACAGACTGTGTGGTCCCGACACATGTGATCTT 542
QY 991 GGCATGCTGTATGGGTACCATGTGTACCCCAAGAAAGCTGCTGTGATTTGAATGAT 1050
Db 543 GGGATGCTGATTTTGAACCTGTGTGATCCGAGCAGAAAGCTGCTCGCTCATAGAAGAT 602
QY 1051 GATGGCTTCCATCAGCCTTCCACACTGCCACAGCTGGGCCACGTTTCAACATGCC 1110
Db 603 GATGGTTTACAAGCTGCTTCCACACAGCCCATGAATTAGGCCACGTTTAAACATGCCA 662
QY 1111 CATGACAATGTGAAAGTCTGTGAGAGGTGTTTGGGAAGCTTCCGAGCCACACCATGATG 1170
Db 663 CATGATGATGCAAGCAGCTGTGCCAGCCTTAATGTGTGACACAGGATTCACCATGATG 722
QY 1171 TCCCGGACCCCTACCATCGACGTGCCAACCCCTGGTGCAGCTGAGTGTGCTATC 1230
Db 723 GCGTCAATGCTTTCCAACTGGACACAGCAGCCTTGGTCTCCTTGCAGTGCCTAGATG 782
QY 1231 ATCCAGGACTTCTGGACAGCGGCACGCTGCTGCTCTGACCAACACCCAGCAAGCCC 1290
Db 783 ATTACATCATTTCTGGATAATGGTCTGCGGAATGTTGTGAGACAGCCTCAAAATCCC 842
QY 1291 ATCTCCCTGCCGAGGATCTCCGGGCGCCAGCTACACCTTGAGCCAGAGTGCAGAGTG 1350
Db 843 ATACAGCTCCCGAGGGGATCTCCCTGGCACCTCGTACGATGCCAACCGGACAGTCCAGTT 902
QY 1351 GCTTTTGGGTGGGTCCAAAGCCCTGCTTACATGCAAGTAC ---TGCACCAAGCTGTGG 1407
Db 903 ACATTTGGGAGGACTCCAAACACTGCCCCGATCGACGACGACATGTAGACCTTTGTGG 962
QY 1408 TGCACCGGAAGGCCAAGGGACAGATGTGTCCAGACGCCGCCACTTCCCTGGGCGGAT 1467
Db 963 TGTACCGGACCTCTGGTGGGTGCTGTGTGTCAACCAACACACTTCCCGTGGCGGAT 1022
QY 1468 GGCACACGCTGTGGCAGGGCAAGCTCTGCTCAAGGGGCTGCGTGGAGAGACAAAC 1527
Db 1023 GGCACACGCTGTGGAGAAAGGAAATGGTGTATCAACGGCAAGTGTGTGAACAAACGAC 1082
QY 1528 CTCACAAGCACAGG ---GTGGATGTTCTTGGGCCAATGGATCCCTATGSCCCC 1581
Db 1083 AGAAAGCATTTTGTATACGCTTTTTCATGGAAGCTGGGAATGTGGGGCCCTTGGGAGAC 1142
QY 1582 TGCTCGCGCACATGTGTGGGGGCTGAGCTGGCGAGGAGGAGTGCACCAACCCACC 1641
Db 1143 TGTTCGAGACCTGCGGTGGAGAGTCCAGTACACGATGAGGAATGTGAACACCCAGTC 1202
QY 1642 CCGTCCAAAGGGGCAAGTACTGCGAGGAGTGAAGGTGAATACCGATCTGCAATCTG 1701
Db 1203 CCAAGAAATGGAGGAAGTACTGTGAAGCAAAACGAGTGCCTACAGATCTCTTAACCTT 1262
QY 1702 GAGCCTGCCCGAGCTCCGGAAGAGTTCGCGGAGGAGCAGTGTGAGGCTTTC 1761
Db 1263 GAGGACTGTCCAGAC ---AATATGGAAGAAACCTTTAGAGAGGAACAATGTGAAGCAC 1319
QY 1762 AAGCGCTACAACACACGACCAACCGGCTCACTCTGCGCGTGGCATGGGTGCCCAAGTAC 1821
Db 1320 AAGGATTTTCAAAAGCTTCTTTGGAGTGGCCCTGCGGTGAATGGATTCCTCCCAAGTAC 1379
QY 1822 TCCGGCGTGTCTCCCGGAGCAAGTGCAGCTCATCTCCGAGGCAATGGCAGTGGCTAC 1881
Db 1380 GCTGGCGTCTCAACAAAGGACAGGTGCAAGCTCATCTGCCAAGCCAAAGGCAATGGCTAC 1439

```
QY 1882 TTCTATGTCTGGCACCACCAAGTGTGGACGGCACCGTGTGCTCTCTCTGACTCCACCTCC 1941
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1440 TTCTTCGTTTGGACGCCCAAGTTGTAGATGTTACTCCATGATAGCCAGATTCCACCTCT 1499
QY 1942 GTCTGTGTCCAAAGCAAGTGCATCAAGCTGGCTGTGATGGGAACCTGGGCTCCAAAG 2001
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1500 GTCTGTGTGCAAGGACAGTGTGTAAGCTGGTTGTGATCGCATATAGACTCCAAAAG 1559
QY 2002 AGATTGCAACAAGTGTGGGTGTGGGGAGACAATAAGAGCTCAAGAAGTGCACCTGA 2061
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1560 AAGTTTGATAAATGTGGTGTGGGGGAAATGGATCTACTTGTAAAAAATATCAGGA 1619
QY 2062 CTCCTTCCACCAAGCCATCATGGGTACAAATTCCTGTGTGTCATCCCGCAGCGCCTCA 2121
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1620 TCAGTTACTAGTGAACAACTGGATATCATATCATCATCAAAATCCAAATGGAGCCAC 1679
QY 2122 AGCATCGACATCCGCCAGCGGTTTACAAAGGCTGATCGGGAT-CACAACATACCTGAC 2180
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1680 AACATCGAAGTGAACAGCGGAACAGAGGGGATCCAGGAACAATGGCAGCTTCTTTC 1739
QY 2181 TCTGAAGAACAGCCAAAGCAAGTACCTGCTCAACGGGCAATTCGTGTGTGTCGGGGTGA 2240
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1740 CATCAAGCTCTGATGSCATATATCTTAATGTGTGACTACATTTGTCACCTTAGA 1799
QY 2241 CGGGACCTGTGTGTAAGGCACTGTCTGCGGTACAGCGGCACGGGCACAGCGGTGA 2300
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1800 GCAAGACATTATGACAAAGGTGTGCTTGAGGTACAGCGGCTCCTCTCGGCGCATGA 1859
QY 2301 GAGCTGCGAGCTTCCCGGCCATCTGGAGCGCTGACCGTGGAGTCTCTCCGTGGG 2360
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1860 AAGATTCCGAGCTTAGCCCTCTCAAGAGCCCTTGACCATCGAGCTTCTTACTGTGG 1919
QY 2361 GAAGATGACACCGCCCGGGTCCGCTACTCTCT 2395
||||| ||| ||||| ||| ||| ||| |||
Db 1920 CAATGCCCTTCGACCTAAATTAATACACCTACT 1954'
```

RESULT 4

```
US-09-392-184-7/c
; Sequence 7, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PROTEASE HOMOLOGS
; FILE REFERENCE: 5800-55
; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 3126
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3126)
; OTHER INFORMATION: reprotolysin (ADAM family of metalloprotease)
US-09-392-184-7
```

```
Query Match 21.0%; Score 598.2; DB 4; Length 3126;
Best Local Similarity 57.4%; Pred. No. 2.4e-125;
Matches 1303; Conservative 0; Mismatches 908; Indels 60; Gaps 10;

QY 157 CTCATTTTTCAGATCAGACATTCAGGAGGACTTTTACTACACCTGACCGCGGTATGT 216
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3122 CTGCGCTCCACCTGTCCGCTTCGGCAAGGGCTTCGTGTGCGCTGG-GGCCGACGAC 3063
QY 217 CAGTCTTGTGCTCCGCTTCTCCACTGAGCATCTGGGGCTCCCGCTCCAGGGGCTCAC 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3062 AGCTTCTGCGCGCCGACTTCAAGATCGAGCGCTCGGGGGCTCCCGCGGGG---ACC 3006
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 GGGGGCTCTTCAGACCTGCGACCGCTGCTTCTATCTGGGGACGTGAACG-CGAGCGG 336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 3005 GGGGCGAGCGGGGGTGCAGCGCTCTCTTCTCCGGGACCGTCCACCGCTCCGCGGAG 2946
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 337 TCCTTCGCTGCTGTGAGCCTGTGCGGGGGCTCCCGGAGGCTTTTGGCTACGAGCGCGC 396
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2945 TCCTTCGCGCGCTGACGCTGTGCGCGGGCTGAGCGGCTCTCTCTGCTGAGCGGCGAG 2886
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 397 GAGTATGTATAGCCCGCTGCCAATGCTAGGCGCGCGCGCGCGGAGCGCAACACCGAG 456
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2885 GAGTTCACCATCCAGCGCAGGCGCGGGGGCTCCCTGGGTGAGCGCACCGCTTCAG 2826
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 GCGCACACCTTCTCCAGCGCGGGGTGTTCCGGGGGGCTTCCGGAGACCCACCTCT 516
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2825 CGCTGGGTCCCGCGGAGCGCCGCCCTCCCGGAGAGCCGAGTGGGAGGTGGAGAGC 2766
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 CGCTCGGGGTG-GCCTCGGGGTGGAACCCCGCCATCTTACGGGGCTTGGACCTTACAA 575
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2765 GGAGAGGGTTCAGAGCGGAGGAGAGGAGACCCAGGAGGAGAGGAGAGAGAGCCAA 2706
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 576 GCGCGCGGGGGGCTTCGGGGAGAGTCTAGCCGGCGCAGGTCTGGGGG-G-CC-CC 630
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2705 GAAGAGGAGCAGAAGCGCTAGCGAGCGCCACCGCCCTGGGGGCCACGAGTAGGACC 2646
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 631 AAGCGTTTCTGTCTATCCCGGGTACGTGGAGACGCTGTGTGTCGCGACGAGTCAATG 690
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2645 AAGCGGTTTGTGTCTGAGCGGCTTTCGTGGAGACGCTGCTGTGGCCGATCGTCCATG 2586
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 691 GTCAGTTTCCAGCGCGCGACCTGGAAACATTTCTGCTGACGCTGCTGCAACGCGGG 750
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2585 GCTGCTTCTACGGGGCGACCTGTCAGAACACCATCTGACGTTAATGCTGTGGCAGCC 2526
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 751 CGACTCTACCGCATCCCGAGCATCTCAACCCCATCAACATCGTTGTGCTCAAGTGTG 810
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2525 CGNATCTAAGCACCCCGACATCAAGAATTCATCAACCTGATGGTGTGTAAGGTGTG 2466
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 811 CTTCTTAGAGATCGTGACTCCGGGGCCCAAGTCAACGGCAATCGCGCCCTGACGTGCGC 870
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2465 ATCGTAGAAGATGAAAAATGGGCCCCAGAGGTGTCGACAAATGGGGGCTTACACTGCT 2406
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 871 AACTTCTGTGCTGCGCAGAGAAGAGTGAACAAGTGAAGTGAAGACACCCCGAGTATGG 930
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2405 AACTTCTGAACATGGCGCGGCTTTTCAACGAGCCGACGCGCCGCCACCCAGAGACATC 2346
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 931 GACACTGCGCATCTCTTACAGGAGGACCTGTGTGG---AGCCACACCTGTGACACC 987
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2345 GACACGGCCATCTGCTCACAGACAGAACTTCTGTGGCAGAGGGGCTGTGTGACACC 2286
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 CTGGGCAATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAGTGTCTCTCATTTAG 1047
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2285 CTGGGTGTGGCAGACATCGGACCATTTGTGACCCCAACAAAAGCTGCTCCGTGATCGAG 2226
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1048 GAGGATGGGCTTCCATCAGCCTTACCACCTGCCAGAGCTGGGCCAGCTGTTCACATG 1107
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2225 GATGAGGGCTCCAGCGGCCCAACCCCTGGCCCATGAATAGAGGACGCTCTCCACATG 2166
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1108 CCCCATGACAATGTGAAGTCTGTGAGGAGGTGTTTGGGAAGCTCCGAGCCCAACACATG 1167
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2165 CCCCAGAGGACTCCAAGCCCTGTACACGGCTCTTCGGGGCCCATGGGAAGCACCACGCTG 2106
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1168 ATGTCCCGACCTCATCCAGATCGACCTGCCAACCCCTGGTCAAGCTGTGAGTGTGCC 1227
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2105 ATGGCACCGCTTCTGCTCCACCTGAACCCAGAGGCTGCTGTCCTCCCTCGAGCGCTATG 2046
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1228 ATCATCAGGACTTCTTGACAGCGGGCAGGCTGACTGCTCTGAGACACCCACAGCAAG 1287
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2045 TATCTCAGAGAGCTTCTTGGAGGGGGGAC-----CGATTTCAGTAT 2003
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1288 CCAATCTCCCTGCCGAGGATCTGCCGGGCGCCAGCTACACCTGAGCGACAGTGTGAG 1347
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2002 TTATGGCCCTGTGTAGTGTAGGTGTAGAGGTGTCATT-----ATGCTCTTTGG 1955
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1348 CTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCTTACATGAGTACTGCAACCAAGCTGG 1407
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1954 CCGGATTTCGGCCACTGCCCCAACACCTCTGCTCAGGAGCTCTCGCCCCAGCTTTGGTGC 1895
Qy 1408 TGCACCGGGAAGGCCAAGGACAGATGGTGTGCCAGCCGCCACTTCCCTCGGGCCGAT 1467
Db 1894 CACACTGATGGGGCTGAGCCCTGTGCCACAGAAAGTGGACGCTGCCCTGGGGCTGAC 1835
Qy 1468 GGCACCAAGCTGTGGAGGGAAGCAAGCTCTGCTCAAAAGGGCCCTGC-----GTGGAGAG 1520
Db 1834 GGCAGCCGTGGGGCTGGCCACTCTGCTCAGAGGCAAGCTGTCTACTGAGAGGA 1775
Qy 1521 ACACAACCTCAAGAACAGCAGAGGTGATGTTCTTGGGCCAAATGGATCCCTATGGCCC 1580
Db 1774 AGTGGAGAGCCCAAGCCCTGTGTAGATGAGGCTGGGACCGCTGGGACCTGGGGAGA 1715
Qy 1581 CTGCTCCGACATGTGTCGGGGCGTGCAGCTGGCCAGGAGCAGTGCACCAACCCAC 1640
Db 1714 ATGTTCTCGGACCTGTGGAGGAGGTACAGTTTTCACACCGTGTGAGTGAAGACCCGA 1655
Qy 1641 CCTTGCACAAGGGGGCAAGTACTCGGAGGAGTGAGGGTGAATACCGATCCTGCAATCT 1700
Db 1654 GCCTCAGAATGGAGGAAGTACTGCCTGGGTGGAGAGCAAGTACCAGTATGCCACAC 1595
Qy 1701 GGAGCCCTGCCCGACCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTAGGCTTT 1760
Db 1594 GGAGGAATGCC-----CCTGACGGGAAAGCTTTCAGGGAGCAGAGTGTGAGAAGTA 1541
Qy 1761 CAACGGCTACAAACACAGCAGCAACCGGCTCACTCTCGCGTGGCATGGGTGCCCAAGTA 1820
Db 1540 TAATGCTCTACAAATACACTGACATGGAGCGGAATCT---CCTCAGTGGGTCCCAAGTA 1484
Qy 1821 CTCGGCGGTGTCTCCCGGAGCAAGTCAAGCTCACTGTCGGAGCAAGTGGCACTGGCTA 1880
Db 1483 TGCCTGGGGTGTCCCGCGGAGCCGCTCAAGTGTGTTCTGCGGAGCCGCGGGAGGAGCGA 1424
Qy 1881 CTTCTATGTCTGTCACCCAGGTGTGGAGCGACGCTGTCTCTCTCTACTCCACCTC 1940
Db 1423 GTTCAAGTGTTCGAGCCAGGTGATGATGGCACCCCTGTGTGGGCGCAAGAACACTGGC 1364
Qy 1941 CGTCTGTCTCAAGCAAGTGCATCAAGGCTGGCTGTGATGGGAACCTGGGCTTCCAGAA 2000
Db 1363 CATCTGTCTCGTGCCAGTGTGTCAAGGCGGCTGTGACCATGTGTGAGTCTGCTCG 1304
Qy 2001 GAGATTTCAGCAAGTGTGGGTGTGGGGGAGACAAATAGAGCTGCAAGAGTGTACTGG 2060
Db 1303 GAAGCTGGACAAATGCGGGGTGTGTGGGGCAAGGCAACTCTGTGAGGAAGTCTCCGG 1244
Qy 2061 ACTCTTCAACCAAGCCCATGATGCTACAATTTGCTGGTGGCCATCCCGCAGGCGCCTC 2120
Db 1243 GTCCCTACCCCAACCAATATGGCTACATGACATGTCAACCATCCCACTGGTGCCAC 1184
Qy 2121 AAGCATGACATCCCGCAGCGGTTTACAAGGGCTGATCGGGATGACAACTACCTGGC 2180
Db 1183 TAATATTGACCTGAAGCAGCGGAGCCACCGGGTGTGTCAGAAGCATGGGAACCTACCTGGC 1124
Qy 2181 TCTGAGAACCAAGCAAGTACTGCTCAAGGGCAATTTGCTGTGTGTGCTGTGCTGGA 2240
Db 1123 GCTGAAGACGCTGATGGCAGTACCTGCTCAACGGCAACTGTGCCATCTCTGCCATAGA 1064
Qy 2241 GCGGACCTGTGTGAGGAGCAGTCTGCTCGGTACAGGGCAGCGGCACAGCGGTGGA 2300
Db 1063 GCAGGACATCTTGTGAAGGGGACCATCTCTGAAGTACAGCGGCTCCATCCCACTCCG 1004
Qy 2301 GAGCCTGAGGCTTCCCGGCGCCATCTCTGGAGCCGCTGACCGTGGAGTCTCTCCGT--- 2357
Db 1003 GCGCTTGCAGAGCTTCCGGCCCTTGGCAGAGCCCTTGACAGTGCAGCTCTCTGACAGTCCC 944
Qy 2358 GGGGAAGTACACCGCCCGGGTCCGCTACTCCTCTCTATCTGCCCAAGA 2408
Db 943 TGGCGAGTCTTCCCGGCAAAAGTCAAAATACACCTCTTTTGTCTCTAATGA 893

RESULT 5

US-09-369-364A-8

; Sequence 8, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hirskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver: 2.1
; SEQ ID NO 8
; LENGTH: 3638
; TYPE: DNA
; ORGANISM: Mus musculus ADAMTS-8
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (278)..(2992)
; NAME/KEY: misc_feature
; LOCATION: (3636)
; OTHER INFORMATION: n = T
US-09-369-364A-8

Query Match 20.7%; Score 590.4; DB 4; Length 3638;
Best Local Similarity 57.3%; Pred. No. 1.4e-123;
Matches 1334; Conservative 0; Mismatches 906; Indels 90; Gaps 11;
Qy 157 CTCATTTTTCAGATCACAGCATTTTCAGGAGGACTTTTACCTACACCTGACCGCCGATGCT 216
Db 440 CTCGGCTTCCACCTCTCGCCCTTCGGCCAGGGCTTCGTGCTGGCCTGGCGCTGACGCC 499
Qy 217 CAGTTCCTTGGCTCCGCCCTTCTCCACTGAGCATCTGGCGTCCGCCCTCCAGGGCTCAC 276
Db 500 AGCTTCCTGGCGCGCAATTCAGATCGAGCGCTTCGGGGGCTCGAGCGCGCG---GCC 556
Qy 277 GGGGCTCTTCAGACTCGGAGCTGCTTATCTCTGGGAGCTGAACCGCGAGCGCGAC 336
Db 557 GGGGCGAGCGGGGACTCGGTGGCTTCTTCTCTGGCACAGTGAATGGAGACGGGAG 616
Qy 337 TCGTTCGTGCTGTGAGCCTGTGCGGGGGGCTTCCCGGAGGCTTTTGGCTACCGAGGCC 396
Db 617 TCGTGGCGGATGAGTGTGTCGCGGCTGGAGCGCTGCTTCTTCTGGCAGCGAG 676
Qy 397 GAGTATGCTATAGCCCGTCCCAATGTAGCGCGCGCGGCGCGCCAGCAAGCCAG 456
Db 677 GAGTTTCAACATCAGCCACAGGGCGCTGGGACTCCCTGGACAGCCTCATCGCCTGAG 736
Qy 457 GCGCACACCTTCTCAGCGCGGGGTGTTCGGCGGGGCTTCCGGGAGAGTCTGGGAGAG 512
Db 737 CGCTGGGGCGGGAGACGGCGCGCAAGACCCCGGCTCGTGGCGCGAAGTTTTCGCC 796
Qy 513 CTCTCGCTCGGGGTGGCTCGGGGTGGAACCCCGCCATCTCTACGGGCCCTGGAGCCCTTA 572
Db 797 CTCCCTCAAGGACTGAGTGGGAGGTGGATGGTAAATGGGCGAGGAGACAGAGAGT 856
Qy 573 CAAGCGCGGGCGGGCTTCGGGAGAGTGTAGCGCGGGCGAGTCTGGG----- 624
Db 857 GACAAGAGAGGACAAGAACAGGAGGGGTGTCTCAAGAGACAGAGACTCC 916
Qy 625 -----CGCGCAAGCGTTTCTGTATCCCG 651
Db 917 CGCAAGTCCCAACCCCTTCGGATCCAAACTAGAACAGAGGTTTGTGTCGAGGCT 976
Qy 652 CGGTAGTGGAGACGCTGGTGTTCGGGACAGAGTCAATGGTCAAGTTCACCGCGCGGAC 711
Db 977 CGCTCGTGAACAACCTTCTGGTGTGATGCTCATGGCTGCTTCTATGGGACCGAC 1036
Qy 712 CTGGAACATATCTGCTGACGCTGTGCAAGCGCGCGGCTCTACGGCATCCCGAC 771
Db 1037 CTGCAAGACCCATCTCAGGTGATGTAATGGGAGCGCGGTAATCAAGACCCCGAGC 1096
Qy 772 ATCCTCAACCCCATCAACATCTGTGGTCAAGGTGCTGCTTCTTAGAGATCGTACTCC 831

```
Db 1097 ATCAGGAACCTCGTCAACCTTGTGTGTAAGTCTAATAGTGGAAAAAGAAATGG 1156
QY 832 GGGCCCAAGGTCAACGGCAATCGCCCTGAGCGTGCACAACTTGTGTGCTGGCAGAAAG 891
Db 1157 GGGCCCGAAGTGTCCGACAACGGGGGCTCACACTGCGCAACTTCTCGAGCTGGCAACGG 1216
QY 892 AAGCTGAACAAAGTGAAGTCAACAGCAACCCGAGTACTGGGACACTGCCATCTCTTACC 951
Db 1217 CGTTTCAACAAAGCCAGCGGACCGCCAGCGGACACTATGACACTGCCATCTTGTACC 1276
QY 952 AGCGAGGACCTGTGTGG---AGCCACCACTGTGTACACCTGGGCGATGGCTGATGTGGT 1008
Db 1277 AGACAGAATCTGTGGGAGGAGAGCAGTGTGACACCTTGGGATGGCAGAGCTTGGC 1336
QY 1009 ACCATGTGTGACCCCAAGAGAAGTGTCTGTCAATGTAGGACGATGGCTTCCATCAGCC 1068
Db 1337 ACCATCTGTGACCCGACAAGAGTGTCTCAGTGATCAAGGATGAGGGACTCCAGCGACC 1396
QY 1069 TTCACCACTGCCACAGACTGGGCCACGTGTTCACATGTGCCCATGACAAATGTGAAGTC 1128
Db 1397 TACACCCCTGGCCCATGAGCTAGGCGACGTTCTCAGCATGCCCCATGATGATCTAAGCC 1456
QY 1129 TGTGAGGAGGTGTTTGGGAAGCTCCGAGCCAAACACATGATGTCGCCGACCTCATCCAG 1188
Db 1457 TGTGTGAGATGTTTGGGCCCATGGGCAAGTACCACATGATGGGCCATCTTTCATCCAC 1516
QY 1189 ATCGACCTGTGCAACCCCTGTGACCTGTGAGTGTGTCATCATCACCAGACTTCTCTGGAC 1248
Db 1517 GTGAACAAGAGCGTCCCTGTCTCCCTGACGTGTCTACCTCACAGAGCTCTCTGGAT 1576
QY 1249 AGCGGGCAGCGTACTGCTCTTGACCAACCCAGCAAGCCCATCTCCCTGCCCGAGGAT 1308
Db 1577 GATGCTACAGGAGATGTCTTCTGATGCCGCCACCTCGGTTCTGCCCCCTCCACAGGC 1636
QY 1309 CTGCGCGGCGGACG-----CTACACCTGAGCCAGCAGTGCAGCTGGCTTTTGGCGTG 1362
Db 1637 CTCCGGGCGACACCCCTCTAGAGCTGGACAGCAGTGAAGCAGATCTTTGGGCGCT 1696
QY 1363 GGCTCCAAGCCCTGTCTTTACA-TGCAGTACTGCACCAAGTGTGTGTCACCGGGAAGGC 1421
Db 1697 GATTTCGGACACTGCCCAACACCTCTGTGGAGGACATCTGTGTCCAGCTCTGTGCCGT 1756
QY 1422 CAAGGGACAGATGGTGTCC-----AGACCCGCCACTTCCCTCGGGCCGAT 1467
Db 1757 CATCGGATAGTATGATGAGCCCATTTGCCACAAAGAAATGTAGCCTGCTCTGGCTGAT 1816
QY 1468 GGCACAGCTGTGGCGAGGGCAAGCTCTGCTCAAGGGGCTCGCTGGGAGGAGACACAAC 1527
Db 1817 GGTACACCTGTGGCCCTGGCACCTGTGCTGTGATGGTAGCTGTGTACTCAAGGAGAT 1876
QY 1528 CTCACAAAGACA-----GGGTGGATGGTTCCTGAGGCCAAATGGGATCCCTATGGCCCC 1581
Db 1877 GTGGAGATCCCAAGGCTGTGTAGATGGAGACTGGGTCTCTGGAACCTGGGGACAA 1936
QY 1582 TGCTCGGCACATGTGTGGGGCGTGCAGCTGCCAGGAGGAGTGCACCAACCCCAACC 1641
Db 1937 TGTCTCGACCTGTGGTGGAGGATACAATTTCTGAACCGTGAATGTGATAATCCAAATG 1996
QY 1642 CCTGCCAACGGGGCAAGTACTGCGAGGGAGTGAAGGTGAATACCGATCTCTCAATCTG 1701
Db 1997 CCTCAGAATGGAGGAAGTTTGTCTGGTGAAGAGATCAAGTACCAATATGCAACACA 2056
QY 1702 GAGCCCTGCCCCAGCTCAGCCTCGGAAAGAGCTTCCGGGAGGAGCAGTGTGAGGCTTTC 1761
Db 2057 GAGGAATGTC-----CACCAAGCGAAAGAAAGCTTCCGGGAGCAGCAGTGTGAGAAATAT 2110
QY 1762 AACGGCTACACCAACAGCACCACCGCTCACTCTCCCTGGCATGGGTGGCCAAAGTAC 1821
Db 2111 AATGCCATACACCA---CACTGACCTGGATGGGAATTTCTGCGATGGGTGCCCAAGTAT 2167
QY 1822 TCCGGCGTGTCTCCCGGGGACAAGTGCAGCTCATCTGCGCGCAATGCGCACTGGCTAC 1881
```

```
Db 2168 TCAGGACTGTCCCCCGGAGACCGATGCAAGCTGTTTTCAGAGCCCGTGGGAGTAUTGAG 2227
QY 1882 TTCTATGTGTGGCACCAAGGTGGTGGACGGCAGCTGTGCTCTCTGACTCCATCTCC 1941
Db 2228 TTCAAAGTGTGTAAGCTAAGTGTATGATGGCACTCTGTGTGACCGGATACTCTGTCC 2287
QY 1942 GTCTGTGTCCAAAGCAAGTGCATCAAGCTGTGTGATGGGAACCTGGGCTCCAAGAAG 2001
Db 2288 ATCTGCTCGGGGGCAATGTGTAAGCTGGCTGTGACCATGTGTGTAACCTCACCTAAG 2347
QY 2002 AGATTCACAAAGTGTGGGTGTGTGGGGGAGACAATAAGAGCTGCAAGAGGTGATGGA 2061
Db 2348 AAGCTGGAATAATGTGGGTGTGTGGGGGCAAAAGGCATCCCTGTAGGAAGATCTCCGT 2407
QY 2062 CTCTTCAACCAAGCCATGCTACAAATTTCTGTGTGGCCATCCCGCAGCGGCTCA 2121
Db 2408 TCITTCACCCCTTCAGTTATGCTACATGACATTTGTACCATCCAGCTGTGTGCCACA 2467
QY 2122 AGCATGACATTCGCCAGCGGTTTAAAGGGCTGATCGGGGATGACAACTACCTTCT 2181
Db 2468 AACATTGATGTAACAGCGGAGTCAACCCAGGGTCAAGGAACGACGCGAGCTACCTGGCG 2527
QY 2182 CTGAAGACAGCCAAAGCAAGTACCTGCTCAACGGGCAATTTCTGTGTGCGGCGTGGAG 2241
Db 2528 CTGAAGACAGCCAAATGGGCAGTACCTGCTCAATGTAACCTGGCCATCTCTGCCATAGAG 2587
QY 2242 CGGCACTGTGTGAAGGGCAGTGTGCTGCTGACAGGGCAGCGGACAGCGGTGGAG 2301
Db 2588 CAAGACATCTGTGTGAAGGGACCATCTCTGAAGTACAGTGGCTCCATGGGTACCTTGGAG 2647
QY 2302 AGCCTGAGCGTTCGCCGGCCATCTCTGGAGCCCTGACCGTGGAGGTCTCTTCCCTG--- 2358
Db 2648 CGGCTGAGAGCTTCCAGGCCCTGCTGAGCCCTTTACAGTACAGCTCCTGACTGTGTCT 2707
...QY 2359 GGAAGATGACACGCGCCCGGCTCCGCTACTCTTCTATCTGCCCAAGA 2408
Db 2708 GGTGAGGTCTTCCTCCAAAGTCAGATATACCTTCTTTGTCCCAATGA 2757

RESULT 6
US-09-122-126B-1
; Sequence 1, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 4192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (406)..(2916)
US-09-122-126B-1

Query Match 20.4%; Score 583.2; DB 4; Length 4192;
Best Local Similarity 60.4%; Pred. No. 6e-122;
Matches 1093; Conservative 0; Mismatches 693; Indels 24; Gaps 7;
```


Db	1134	AACAGTGTGGCAGCAGCAGCCAAAGCGCTTCAAGCACCAAGCATCCGCAATCCTGTGCAG	11193
Qy	789	CATCGTTGTGGTCAAGTGTCTCTTAGAGATCGTGACTCCGGGCCCAAGTCAACGG	848
Db	1194	CTTGTGTGTGATCTCGGCTAGTAGTCTGGGTCTCAGCGAGAGAGGGGCCCCAAGTGGGGCC	1253
Qy	849	CAATCGGCGCTGAGCGCTCGCAACTTCTGTCCCTGGCAGAGAAGCTGAACAAAGTGAG	908
Db	1254	CAGTGTGCCAGACCCCTGGCAGCTTCTGTGCCTGGCAGGGGGCCTCAACACCCTGA	1313
Qy	909	TGACAAGACACCCGAGTACTGGGACACTTGCCATCTCTTACACAGGACAGACCTGTGTGG	968
Db	1314	GGACTCGGACCTGACCACTTTGACACAGCCATTCTCTTTACCCGCTCAGACCGTGTGG	1373
Qy	969	AGCCACCACTGTGACACCCCTGGCATGGCTGATGTGGTACCATGTGTGACCCCAAGAG	1028
Db	1374	AGTCTCCACTTGGACACGCTGGGTATGGCTGATGTGGCACCCTCTGTGACCCGGCTCG	1433
Qy	1029	AACTGCTCTGTCAATTGAGGACGATGGGCTTCATCAGCTTTCACACTGCCACAGACT	1088
Db	1434	GAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAGCCTTCACTGCTCTCATGAAC	1493
Qy	1089	GGCCACGCTGTTCAACATGCCCATGACAATGTGAATCTGTGAGGAGTGTGGAA	1148
Db	1494	GGTCAATGCTTCAACTGCTCCATGACAACTCCAAAGCCATGCATCAGTTTGAATGGCC	1553
Qy	1149	GTCCGAGCCAA---CCAGATGATGTCCCGACCCCTCATCAGATPCAGCGTGCACAACC	1205
Db	1554	TTTGAGCACTCTCGGCATGTCTATGGCCCTGTGATGCTCATGTGATCTCTCAGGAGCC	1613
Qy	1206	CTGCTCAGCTCGAGTGTGCCATCATCAGCGACTTCTGTGACAGCGGACGCTGACTG	1265
Db	1614	CTGGTCCCTCTGAGTGTCCCGCTTCATCACTGACTTCTTGACAAATGGCTATGGCACTG	1673
Qy	1266	CTCTCTGGACCAACCCAGAACCCCATCTCCCTGCCGAGGATCTGCCGGCGCCAGCTA	1325
Db	1674	TCCTTTAGACAAACAGAGCTCCATTGCATCTGCTGTGACTTTCCTGTGCAAGGACTA	1733
Qy	1326	CACCTTGACCCAGCTGCGAGCTGGCTTTTGGCTGGGCTCCAGCCCTCTCCCTTACAT	1385
Db	1734	TGATGTGACGGCAGTGGCAGCTGACCTTTCGGGCCCCGACTCAGCCCATGTTCACAGCT	1793
Qy	1386	---GCAGTACTGCACCAAGCTGTGTGCAACGGGAAGCCCAAGGACAGATGTGTGCCA	1442
Db	1794	GCCGCGCCCTGTGCTGCCCTCTGTGTGCTCTGGCCACCTCAATGSCCATGCATGTGCCA	1853
Qy	1443	GACCGGCACCTTCCCTTGGCGCATGTGCACACAGCTGTGGCGAGGGCAAGCTCTGCCTAA	1502
Db	1854	GACCAACACTCCGCTTGGCGCGATGGCACACCTTGGCGGCCCGCACAGGCGCTCATGCG	1913
Qy	1503	AGGGCTCTGGTGGAGAGACAACTCTCAACAAAGCACAGGT-----GGATGGTTCCTG	1556
Db	1914	TGTCGTGCTCCACATGGACAGCTCCAGGACTTCAATATTCACAGGCTGTGGCTG	1973
Qy	1557	GGCCAAATGGATCCCTATGGCCCTGTCTGCGGCACATGTGTGGGGGCGTGCAGCTGGC	1616
Db	1974	GGGTCTTGGGGACCATGGGGTGACTGCTCTCGGACCTGTGGGGTGTGTCCAGTTCTC	2033
Qy	1617	CAGGAGCACTGCACCAACCCACCCCTGCCAAACGGGGCAAGTACTCCGAGGAGGTGAG	1676
Db	2034	CTCCGAGACTGCACGAGGCTGTCCCCGGAAATGTTGGCAGTACTGTGAGGGCCCGCG	2093
Qy	1677	GGTAAATACCGATCTCTGCAATCTGAGGCCCTGCCCGAGCTCAGCCTCCGGAAGAGCTT	1736
Db	2094	TACCGCTTCCGCTCTGCAACACTGAGGACTGCCAA---CTGCTCAGCCCTGACCTT	2150
Qy	1737	CCGGGAGGACGATGTGAGCTTTCAAGGGCTTCAACACAGCAGCACCAACCGGCTCACTCT	1796
Db	2151	CCGCGAGGAGCATGTGTGCTTACAAACACCGCACCGACCTTTCAGAAGCTTCCCGAGG	2210
Qy	1797	CGCGTGGCATGGGTGCCCAAGTACTTCCGGCGCTGTCTCCCGGGACAAGTCAAGCTCAT	1856

Db	2211	GCCCATGGACTGGGTCTCTCGGTACACAGCGGTGGCCCCCAGGACCGATGTCACAACTCAC	2270
Qy	1857	CTGCGAGCCCAATGGCACTGGCTACTTCTATGCTGTGCACCCCAAGGTGGTGGACGGCAC	1916
Db	2271	CTGCGAGGCCCGGCACTGGGCTACTACTATGCTGTGAGCCACGGGTGGTAGATGGAC	2330
Qy	1917	GCTGTGTTCTCTTGACTCCACCTCCGTTCTGTGTCGAAGCAAGTGCATCAAGGCTGGCTG	1976
Db	2331	CCCTGTTCCCGGACAGCTCCTCGTCTGTCTCCAGGGCGATGCATCCATGCTGGCTG	2390
Qy	1977	TGATGGGAACCTGGGCTCCAAAGAAGATTCACAAAGTGTGGGTGTCTGGGGGAGACAA	2036
Db	2391	TGATCGCATCATTTGGCTCCAAAGAAGTTTGACAAGTGCATGGTGTGCGGAGGGACGG	2450
Qy	2037	TAAGAGTGCAGAAGGTGACTTGGACTCTTCACCAAGCCCATGCATGGCTACAATTTCTGT	2096
Db	2451	TTCTGGTTGCAGCAAGCAGTCAGGCTCTTCAGGAAATTCAGGTACGGATACACAATGT	2510
Qy	2097	GGTGGGCATCCCGCAGGCGCTTCAAGCATGCACATCCGCGCAGCGCGGTTCACAAAGGCT	2156
Db	2511	GGTCACATATCCCGCGGGGCCACCCACATTTCTCGGGCAGCAGGGGAAACCCCTGGCCA	2570
Qy	2157	GATCGGGATGACAACTACTTGGCTGTGAAGAACAGCCAAAGTAGTACCTGCTCAACGG	2216
Db	2571	CCGGAG-----CATCTACTTGGCCCTGAAGCTGCCAGTGGCTCCTATGCCCTCAATGG	2624
Qy	2217	GCATTTCTGGTGTGCGCGGTGGAGCGGACCTGGTGGTG--AAGGGCAGTCTGC-TTCG	2273
Db	2625	TGAATACACCTGATGCCCTCCCCACACATGTGTTACTGCTGGGCAGTCACTGCTGG	2684
Qy	2274	GTACAGCGGCACGGGCACAGCGGTGGAGAGCCTGCAGGCTTCCCGGCCCATCTCTGGAGCC	2333
Db	2685	CTACAGCGGGGCCACTCGAGCCTTCAGAGACACTGTCAAGCCCATGGGCCACTGGCCCAAGCC	2744
Qy	2334	GCTACCGTGGAGTCTCTCTCGTGGGGAAGATGACACGCCGCCGGGTCCGCTACTCTCT	2393
Db	2745	TTTGACACTCGAAGTCTTAGTGTCTGGCAACCCCGAGACACACGCCCTCCGATACAGCTT	2804
Qy	2394	CTATCTGCC	2403
Db	2805	CTTCGTGCC	2814
RESULT 7			
US-09-122-126B-14			
; Sequence 14, Application US/09122126B			
; Patent No. 6451575			
; GENERAL INFORMATION:			
; APPLICANT: Bristol-Myers Squibb Company			
; TITLE OF INVENTION: AGGREGAN DEGRADING METALLO PROTEASES			
; FILE REFERENCE: DM6909			
; CURRENT APPLICATION NUMBER: US/09/122,126B			
; CURRENT FILING DATE: 1998-07-24			
; NUMBER OF SEQ ID NOS: 21			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 14			
; LENGTH: 3250			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (121)..(2910)			
US-09-122-126B-14			

Query Match 19.2%; Score 549; DB 4; Length 3250;
Best Local Similarity 57.8%; Pred. No. 2.7e-114;
Matches 1044: Conservative 0; Mismatches 745; Indels 1

QY 617 GGTCTGGCGCGCCAAAGCGTTTCGTGCTATCCCGCGGTACGTGGAGACGCTGGTGGTCG 676
db 884 GTTGGCGCGCGCGCGCGCTCCATCTCCGGGCGCCAGCGTGGAGCTGCTTCTGGTGG 943
QY 677 CGGACGAGTCAATGGTCAAGTTCCACGCGCGAGCTGGAAACATTATCTGCTCACGCTGC 736

LOCATION: (2)...(2623)
US-09-369-364A-14

Query Match 16.5%; Score 469.8; DB 4; Length 2625;
Best Local Similarity 56.4%; Pred. No. 1.5e-96;
Matches 1008; Conservative 0; Mismatches 752; Indels 27; Gaps 6;
QY 629 CCAAGCGTTTCGTCTATCCCGGGTACGTGGAGACGCTGGTGGTCGCGACGAGTCAA 688
DB 372 CCAAAACGCTTTCGTCTACCCACGCGTTGTAGAGGTGATGGTGGTCGACACAGGA 431
QY 689 TGGTCAAGTTCCACGGCGGGAGCTGGACATATCTCTGAGCGTGTGTCGACAGCGGG 748
DB 432 TGGTTTATACACCGGAGCAACCTTCAACATATATCTTAACCTTAAATCTCCATTGTAG 491
QY 749 CGCGACTCTACCGGCATCCAGCATCTTCAACCCCATCAACATGCTTGTGTCAGGTGC 808
DB 492 CTCTCTATATAAGACTCAAGTATGGAAATTTAAATTAATTTATTTGTAACCTTAG 551
QY 809 TCGTCTTTAGAGATCGTGACTCCGGGCCCAAGGTCACCGCAATTCGCGCCCTGACGCTGC 868
DB 552 TTGTGATTCATAATGAACAGGAAGACCTTACATAAATTTCAAAGCCAGACACATTA 611
QY 869 GCAACTTCTGCTGCGGAGAGAGCTGAACAAGTGAAGTGAAGACACCCGAGTACT 928
DB 612 AGAAGCTTTGGCAGTGGCAGCACTCAAAAGAACTACTTGGGTGGATTTCAGCAC- 664
QY 929 GGGACACTGCCATCTCTTCCAGGAGGACCTGTGTGGAGC--CACCACCTGTGACA 985
DB 665 --GACACAGCGTCTGTGTCACAGGGAAGATCTGCAGAGCTCAGGACAAATGTGACA 722
QY 986 CCCTGGGATGGCTGATGTGGTACCAATGTCGACCCCAAGAGAGTGTCTGTCTATTG 1045
DB 723 COTTAGGCTTGTGCTGACAGGGAACCATTTGCGACCCCTACCGAAGCTGTTCATTAGT 782
QY 1046 AGGAGATGGCTTCCATCAGCTTCAACCTGCCCAGAGCTGGGCAAGTGTTCACAA 1105
DB 783 AAGAGTGGCTGAGCACAGCTTTCACATAGCTCAGAGCTGGGCCATGTGTTTAATA 842
QY 1106 TGCCCCATGACATGTGAAAGTCTGTGAGGAGTGTGTTGGAAAGCTCCGAGCCCAACACA 1165
DB 843 TGCTTCAGATGACAGCAATAAATGCAAGGAAGAGGATTAAG--AGTCCCAGCATG 899
QY 1166 TGATGTCCCGACCTCTATCAGATGACCGTCCCAACCCCTGTGAGCTGACGCTG 1225
DB 900 TCATGTGACCAACACTGAAGTCTTACACCAACCCCTGGATGTGTCGTCAGTCCGA 959
QY 1226 CCATCATCAGCGACTTCTGTGACAGCGGACGCTGCTGCTGACCAAG--CCA 1282
DB 960 AATACATCAGTGTCTTAGACACTGGGTACGAGAGTGTGCTGAAAGCACTCGAT 1019
QY 1283 GCAAGCCCATCTCCCTGCGGAGGATCTGCGGGCGCCAGCTACACCTGAGCCAGCAGT 1342
DB 1020 CCAGGACCTATCTTGTGCTTCCCACTGCGCGCCCTTCTTACAAAGTCAATAAACAAT 1079
QY 1343 CGAGCTGGCTTTTGGCTGGGCTCCAGCCCTCTCTATCATGCACTGTGCACCAAGC 1402
DB 1080 GTGAAGTATTTTGGGCCAGGCTCTCAAGTGTGCGCCCTATATGATGAGTGCAGAGCGG 1139
QY 1403 TGTGTGACCGGGAAGGCCAAGGACAGATGTTGTCGACAGCCGCACTTCCCTGGG 1462
DB 1140 TCTGGTGAATAATGTGATGGAGCACACAAGGCTGCAAGCTGCAAGCTCAATAAACAAT 1079
QY 1463 CCNATGGACAGCTGTGGCAGGGCAAGCTCTGCTTCAAGGGGCGCTGCTGGAGAGAC 1522
DB 1200 CAGATGGAACCGAGTGTGAGCTGGAAGCACTGCAAGTTGGATTGTGTTCCCAAG 1259
QY 1523 ACACTTCAACAACAGCAGGTGATGTTCTCTGGGCGCAATGGATTCCTATGGCCCT 1582
DB 1260 AAA--TGGAGGGCCCTGCAATGTGATGATCTCTGGGAGGTGTGAGCCACTTTGGGACCT 1316
QY 1583 GCTCGGCACATGTGTGGGGCTGTGAGCTGCGGAGGAGGCTGCAACCAACCCACCC 1642

DB 1317 GCTCAAGAACGTGTGGAGGAGGATCAAAACAGACCCATCAGAGTGCACAGACACG 1376
QY 1643 CTGCCAACGGGGCAAGTACTGCGAGGAGGTGAATACCCGATCCTGCAATCTGG 1702
DB 1377 CAAAAAATGTTGGGAAGTACTGTGTAGGAAGAGAAATGAAGTTCAATCCTGCAACACGG 1436
QY 1703 AGCCCTGCCAGCTCAGCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGAGCCTTCA 1762
DB 1437 AGCCCTGCATGAAGAGCAAGCGGAG-----ACTTCCGAGAGGAGCAGTGTGCTACCTTTG 1490
QY 1763 ACAGCTAAACACAGCACCACCAACCGGCTCACTCTCGCCGTGGCATGGTGCCTCCCAAGTACT 1822
DB 1491 ATGSCAAACACTTCAACATCAATGGTCTGCTGCCAGCGTACCTGGTTTCTTCAAGTACA 1550
QY 1823 CGGCGTGTCTCCCGGGACAAGTGAAGCTCACTCTGCGGAGGCAATGGCACTGGTACT 1882
DB 1551 CGGGAATTTTGTGAAGAGCGGTTGCAAGTTGTCTGACAGTGGCAGGAGCAACACGCT 1610
QY 1883 TCTATGTGCTGGCACCACCAAGGTGTGGAGCGACGCTGTCTCTCTGACTCCACCTCCG 1942
DB 1611 ACTACAGCTCCGAGACAGAGTATTGACGGAACCCCTTGTGCGCAGGACACAAATGACA 1670
QY 1943 TCTGTGTCCAAAGCAAGTGCATCAAGCTGGCTGTGTGATGGAAACCTGGGCTCCCAAGAAGA 2002
DB 1671 TCTGTGTCCAAAGCCTTTGCGCGCAAGCTGGATGTGATCATATTTTAACTCAAGTCC 1730
QY 2003 GATTGCAAGTGTGGGCTGTGTGGGGGAGACAATAGAGCTGCAAGAGGTGACTGGAC 2062
DB 1731 GGAAGATAAATGTGGGATTTGTGGTGGAGATAATTTCTCATGCAAAACAGTGGCAGGAA 1790
QY 2063 TCTTTCACCAAGCCCATGCTGGCTACAATTTCTGTGTGGCCATCCCGCAGGCGCTCAA 2122
DB 1791 CATTTAAGCTGTCCATTTATGTTTACAATCTGTTCCGAATTCGCGTGGTGTGCTACCA 1850
QY 2123 GCATGACATCCCGCAGCGGCTTCAAAAGGCTGTATCGGGGATGACAACTACCTGGCTC 2182
DB 1851 GCATGACGTGCTGACGACAGCTTCTCAGGGAAGTCTCAGGATGACAACTACCTAGCTT 1910
QY 2183 TGAAGAACGCCAACAGTACTGCTCAACGGGCAATTCGTTGGTGTGCGCGGTGGAGC 2242
DB 1911 TATCAACAGTAAGGTGAATTCCTGCTAAATGGAGACTTTGTGTCTCCATGTCCAAA 1970
QY 2243 GAGACCTGGTGTGAAGGCGAGTCTGCTGCGGTGACAGCGCAGCGGACAGCGGTGGAGA 2302
DB 1971 GCGAGTCCGCGTGGGAGCGCGCTCATTTGAGTACAGCGSATCGGACATGTGTGANA 2030
QY 2303 GCTGAGGCTTCCCGGCCCATCTCGAGCGCTGACCGGTGAGGTGCTCTCCGTGGGGA 2362
DB 2031 GACTGAAGTGTACGGACCGTATCGAGGAAGAACTTCTCTCAGGTGTGTCGCGTGGGAA 2090
QY 2363 AGATGACACCGCCCGGCTCCGCTACTCTCTTCTATCTGCCCCAAAGAG 2409
DB 2091 AGCTGTATAACCCAGATGTGCGGTACTCATTAATATTCCTCATTTAG 2137

RESULT 10

US-09-369-364A-12
Sequence 12, Application US/09369364A
Patent No. 6391610
GENERAL INFORMATION:
APPLICANT: Apte, Suneel
APPLICANT: Hurskainen, Tiina L.
APPLICANT: Hirohata, Satoshi
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
FILE REFERENCE: 26473/4007/10-30-00
CURRENT APPLICATION NUMBER: US/09/369,364A
CURRENT FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 5804
TYPE: DNA
ORGANISM: Homo sapiens ADAMTS-9

RESULT 11
US-09-130-491-7
; Sequence 7, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodgearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491

```
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2114
; TYPE: DNA
; ORGANISM: Rattus rattus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1445)
US-09-130-491-7

Query Match      11.1%; Score 318; DB 4; Length 2114;
Best Local Similarity 57.8%; Pred. No. 1.8e-62;
Matches 705; Conservative 0; Mismatches 470; Indels 45; Gaps 6;

QY 1204 CCTCGTTCAGCTGAGTGGCCATCATACCGACTTCCTGGACAGGGGCACGGTGC 1263
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 CCTCGTTCCTCGAGTGGCTGTCTACCTCAGGAGCTCCTGGATGATGGTACGGAGAC 80
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1264 TGCCTCCTGGACCAACCCAGCAAGCCATCTCCCTGCCGAGAGTCTGCCGGGCCAGC 1323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 TGCCTCCTGGATGATGGCCACAGCACCTCTATAGAGCTGGACACAGT----- 129
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1324 TACACCTTGAGCCAGCTGGAGTGGCTTTTGGCGTGGCTCCAGCCCTGCTCTAC 1383
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 -----GCAAGCAGATCTTGGGCTGATTTCCGACACTGCCCAACACCTCTGTG 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1384 ATGCAGTACTGCACCAAGCTGTGTGCACCGGGAAGGCCAAGGACAGATGGTGGCC-- 1441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 GAGGACATCTGTGCCAGCTGTGTGGCGTCATCGGATAGTATGAGCCCATTTGCCAC 239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1442 ---AGACCCGCCACTTCCCTGGGCCGATGGACACAGCTGTGGGAGGCAAGCTTGC 1497
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 ACAAGAATCCAGCTGTCTGGGCTAGGTAGCTAGCCCTGTGGCCCTGGGCACCTTGC 299
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1498 CTCGAAGGGGCTGTGTGGAGAGACACAACCTCAACAGCAC-----GGTGGATGGT 1551
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 CTGGATGTAGCTGTGTGCTCGGGAGGAGTAGAGATCCCAAGGCTGTGTAGATGGA 359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1552 TCTTGGGCCAATAGGATTCCTATGTGCCCTGTCTGCGCACATGTGTGGGGCGTGCAG 1611
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 GACTGGGTCCTCGGGACCTGGGACAAATGTCTCGCACCTGTGTGGAGGATACAG 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1612 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTGGAGGGA 1671
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 TTTTCGAACCTGTAGTGTGATAATCCAGACCTCAGAAATGGAGAAATTTTGCCTGGGA 479
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1672 GTGAGGTGAATACCGATCTGCAATCTGAGGCCCTGCCCCAGCTCAGCTCCGGGAAG 1731
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GAGAGATCAAGTACCAATCTTGCACAGACAGAGAAATGTC-----CACCAATGGAAAA 533
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1732 AGCTTCGGGAGGAGCAGTGTAGGCTTTCAACGGCTTACCAACACACACCAACCGGCTC 1791
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 534 AGCTTCAGGAGGAGCAGTGTGAAATAATAATATGCTTACAAACCA---CAGCGACCTGGAT 590
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1792 ACTCTCGCGGTGGATGGTGGCCCAAGTACTCGCGGTGTCTCCCGGGGACAAAGTCAAG 1851
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 591 GGAATTTCTCTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 650
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1852 CTCATCTCGCGAGCAATGCAGTGGCTACTTCTATGTGTGGCACCCCAAGTGGTGGAC 1911
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 651 CTGTTTTCAGAGGCCCTGGAGAGTGTCAAGTGTGTGAAACTAAGGTGATCGAT 710
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1912 GGCAGCTGTCTCTCTGACTCCACCTCGCTGTGTGTCAAGCAAGTGCATCAAGGCT 1971
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 711 GGCAGCTGTGTGGGACCGGATCTCTTGGCCATCTGTGTGGGGGACAGTCCGTAAAGCT 770
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1972 GGCTGTATGGGAACCTGGGCTCCAAAGAGATTCGACAAGTGTGGGGTGTGTGGGGGA 2031
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 12

```
US-09-369-364A-10
; Sequence 10, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 739
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-8
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(737)
US-09-369-364A-10
```

Query Match 9.1%; Score 259.4; DB 4; Length 739;

Best Local Similarity 64.4%; Pred. No. 1.9e-49;

Matches 405; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

QY 629 CCAAGCGTTTCGTGTCTATCCCGCGGTAGTGGAGACGCTGGTGGCGGACGAGTCAA 688

Db 55 CCAAGCGTTTCGTGTCTATCCCGCGGTTCGTGGAGCGCTGCTGGTGGCGGATGCTCA 114

QY 689 TGGTCAAGTTCACGGCGGGACCTCGGACATTTCTGTGACGCTGTGTGGCAACGGGG 748

Db 115 TGGTGCCTTCTACGGGGCGGACCTCGAAGACCATCTCTGTTAATGTCTGTGGCAG 174

QY 749 CGCGACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCTGTGTGGTCAAGTCC 808

Db 175 CCGGATCTTCAAGCACCCCGCAGCATCAAGATTCATCAACCTGATGGTGTAAAGTGC 234

QY 809 TGTCTTTTAGATGCTGACTCCGGGCCCAAGTCCAGGCAATCGGGCCCTGAGCCCTGC 868

Db 1191 ACCTTCTTCGTCCCAATGA 1210

QY 2389 TCCTTCTATGCCCCAAGA 2408

Db 1191 ACCTTCTTCGTCCCAATGA 1210

QY 2272 CGGTACAGCGGCACGGGCACAGCGGTGGAGAGCTTCAGGCTTCCCGGCCCATCTGGAG 2331

Db 1071 AAGTACAGTGTTCATGGCCACCTCGGAGCGCTGCAGAGCTTCCAAGCCCTCCCTGAG 1130

QY 2332 CCCTGACCGTGGAGTCTCTCCGTG---GGGAAGATGACACCGCCCGGCTCCGCTAC 2388

Db 1131 CCTCTTACAGTACAGCTCCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1190

QY 2389 TCCTTCTATGCCCCAAGA 2408

Db 1191 ACCTTCTTCGTCCCAATGA 1210

Db 235 TGATCGTAGAAGATGAAATATGGGCCAGAGGTGTCGACATGGGGGCTTACACTGC 294
Qy 869 GCACCTCTCTGCTGGCAGCAAGAGCTGACAAAGTGAATGACAGACACCCCGAGTACT 928
Db 295 GTAACTTCTGCAACTGGCAGCGCGTTTCAACAGAGCCCAAGACCCCGAGAGCACT 354
Qy 929 GGGACACTGCATCTCTTCCACAGCAGGACCTGTGTG3---AGCCACACCTGTGTGACA 985
Db 355 ACAGACAGGCGCATCTCTGCTCACCAGACAGAACTTCTGTGGCAGAGGGGCTGTGTGACA 414
Qy 986 CCCTGGGCAATGGTGTATGGTGGTACCATGTGTGACCCCAAGAGAGTCTGCTGTGCAATG 1045
Db 415 CCCTGGGTGGGAGACATCGGAGCAATTTGTGACCCCAACAAAGCTGCTCGGTGATCG 474
Qy 1046 AGGACGATGGCTTCCATGACGCTTCCACACTGCCACAGCTGGCCACAGTGTTCACAA 1105
Db 475 AGGATAGGGGCTCCAGGCGGCCACACCTGTGCCCATGAACATAGGGCACGTCCTCAGCA 534
Qy 1106 TGCCCCATGACAATGTGAAGTCTGTGAGGAGGTGTTTGGAGCTCCGAGGCCAACCA 1165
Db 535 TGCCCCACGACGACTCCAGCCCTGCACACGGCTCTTCGGGCCCATGGCAAGCACACG 594
Qy 1166 TGATGCCCCGACCTCATCCAGATGACCGTGCACCCCTGGTTCAGCTGCAGTGTG 1225
Db 595 TGATGGACCGCTGTGTGTCACCTGAACCCAGACGCTGGCCCTGCTCCCTTGAGCGCCA 654
Qy 1226 CCATCATCAGCGACTTCTCTGGACAGCGG 1254
Db 655 TGTTCTCAGGCTGCCACCTGCAGGGTGG 683

RESULT 13

US-09-392-184-5/c

; Sequence 5, Application US/09392184

; Patent No. 6395889

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN

; FILE REFERENCE: 5800-55

; CURRENT APPLICATION NUMBER: US/09/392,184

; CURRENT FILING DATE: 1999-09-09

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 5357

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(5357)

; OTHER INFORMATION: repolysin (ADAM family of metalloprotease)

; NAME/KEY: misc_feature

; LOCATION: (1)...(5357)

; OTHER INFORMATION: n = A,T,C or G

US-09-392-184-5

Query Match

Best Local Similarity 8.6%; Score 245.6; DB 4; Length 5357;

Matches 926; Conservative 0; Mismatches 779; Indels 83; Gaps 13;

Qy 630 CAACGGTTTGTGTCATCCCGGTACGTGGAGACGCTGGTCCGACAGTCAAT 689
Db 4617 CCAGCGTGGTTCAGCAAGAGAGTGGGTGGAGACCTGGTAGTACGTATGCCAAAAT 4558
Qy 690 GGTCAAGTTCACAGG---CGCGGACCTGGAACTATCTGCTGACGCTGTGGCAACGGC 746
Db 4557 GGTGAGTACACAGCAGCAGCGAGGTGAGAGCTATGCTGACCATCATGACATGGT 4498
Qy 747 GGCAGCTCTACCGCCATCCAGCATCTCAACCCCATCAACATCGTTGTGTCAAGGT 806
Db 4497 GGCTGGCTGTTCATGACCCAGCATTTGGGAACCCCATCCATCATCCATTTGTGGCCT 4438

```
Db 3404 GTACTTTCCGAGAAAGCTGCGGGACGCGCTGGTGGATGGCACCCTGCTACAGGTCCG 3345
QY 1935 CACCTCCG-----TCTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGTAGTGGAACT 1988
Db 3344 AGCCAGCGGGACCTCTGCATCAACGGCATCTGTAAGAAGCTGGGCTGTGACTTTCGAGAT 3285
QY 1989 GGGCTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGACAAATAGAGCTGCAA 2048
Db 3284 TGACTCCGGTGCTATGGAGGACCGCTGTGGTGTGTGCCAGCGCAACGCTCCACCTGCCA 3225
QY 2049 GAAGGTGACTGGACTTTCACAAAGCCCATGCATGGCTACAAATTCGTGGTGGCCATCCC 2108
Db 3224 CACCGTAGCGGGACCTTCGAGGAGCGCGGAGGCTGGGTATGTGGATGTGGGGCTGAT 3165
QY 2109 CGCAGGCGCTCAAGCATCGACATCCGCCAGCGGGTTACAAAGGGTGTATCGGGGATGA 2168
Db 3164 CCCAGCGGGCGCA-----CGCAGATCCGATCCCAAGAGGTTGCGGAGGCTGC 3117
QY 2169 CAACTACTCTGGCTTGAAGAACAGCAAGGCAAGTACCTGCTCAACGGTATTTCTGTGGT 2228
Db 3116 CAACTTCTCGCACTCGGAGTAGGACCCCGGAGAACTACTTCTCAATGTGGCTGGAC 3057
QY 2229 GTCGCGGTGAGCGGGACCTGTGTGTGAAGGCGAGTCTGTGCGGTACAGCGGACGGG 2288
Db 3056 CATCCAGTGAACGGGACTTACCAGGTGGCAGGACCACTTACATACGACGCGAGGG 2997
QY 2289 CACAGCGGTGAGAGCTGTGAGGTTCCCGGCCCATCTTGAGCGCGGTGACCGTGGAGGT 2348
Db 2996 CA---ACTGGGAGAACTCACTGCTCCCGGGTCCCAAGAGGCGCTGTCTGGATCCAGCT 2940
QY 2349 CCTCTCGTGGGGAGATGACACGCGCGCGGTCCGCTACTCTCTTA 2396
Db 2939 GCTGTTCCAGGAGCAACCTGGGGTGCACTAGGAGTACACCATCCA 2892
```

RESULT 14

```
us-09-369-364A-6
; Sequence 6, Application us/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3218
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-7
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(3003)
us-09-369-364A-6
```

```
Query Match 8.4%; Score 238.6; DB 4; Length 3218;
Best Local Similarity 51.3%; Pred. No. 1.3e-44;
Matches 932; Conservative 0; Mismatches 799; Indels 84; Gaps 13;

QY 562 CTGAGCCCTTACAGCGCGGGGGGGCTTCGGGGGAGAGTCTGAGCGCGCGAGGTCT 621
Db 646 CTGAGTCTCGACGGAGGCTTGGAGAGCGGACAGTGGCGGGCCACGGCTGAGG 705
QY 622 GGGCGCGCAAGGTTTCTGTCTATCCCGCGGTAGCTGAGACGCTGTGTGTGCGGAC 681
Db 706 CGTCTACACCGCGGTGCTGACGAAGAAGTGTGTGAGACCGCTGTGTAGTCTGAT 765
QY 682 GAGTCAATGGTCAAGTTCACGCG---CGCGGACCTGGAAACATTATCTGCTGACGCTGTG 738
```

```
Db 766 GCCAAATGTTGAGTACCAGGACAGCCGAGGTTGAGAGCTATGTGTGACCATCATG 825
QY 739 GCAACGCGCGGCGACTCTACCGCCATCCAGCATCTCAACCCCATCAACALCTGTGTG 798
Db 826 AACATGTGTGCTGGCTGTTTATGATACCCAGCATTTGGGAACCCCATCCACATCACT 885
QY 799 GTCAAGGTGCTGCTTCTTAGAGATGTGATCCGGGCGCAAGGTACCGGCAATTCGGGGC 858
Db 886 GTGCGGCTGCTGCTGCTGGAAGATGAGGAGGAGACCTAAAGATCACGCAACCATGCAJAC 945
QY 859 CTGACGCTCGGCAACTTCTGTGCTGCGAGAGAACTCTGAACAAAGTGTGAGTACACAGJAC 918
Db 946 AACACCTCTGAGAGCTCTTGAAGTGGCAGAAAGCATCAACATGAAGGGGATGTCAT 1005
QY 919 CCGAGTACTGGGACACTGCCATCTCTTACACAGCAGGACCTGTGTGGAGCCACCA-- 976
Db 1006 CCCTGCACCATGACACTGCCATCTCTTACACAGAAAGGACCTGTGTGACCCATGAAC 1065
QY 977 ----CCTGTGACACCTCGGCGATGGCTGTATGTGGGTACCATGTGTGACCCCAAGAGAAGC 1032
Db 1066 CGGCCCTGTGAGACCTCGGACTGTCCCATGTGGCGGCGATGTGCCAGCGCCAGCTCGG 1125
QY 1033 TGCTCTGTCTAGGAGCATGGCTTCCATCAGCCTTCCACACTGCCCTGCCCTACAGCTGCGC 1092
Db 1126 TGAGCATCAACGAGGACAGCGGCGCTGCGCTGCGCTTCACTGTAGCCCAACAGCTCGG 1185
QY 1093 CAGCTGTTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGGTGTTTGGGAAGCTC 1152
Db 1186 CACAGTTTGGCATTCAGCATGACGGAAGCGGCAATGACTGTGAGCCGTTGGGAACGA 1245
QY 1153 CGAGCCAAACCATGATGTCCCGACCTTCATCCAGATCGACCGTGCACACCCCTGTGTC 1212
Db 1246 CCTTTTCATC-----ATGTCTCCACAGCTCTGTACGACGCGCTCCCTCACCTGCTCC 1299
QY 1213 GCTGCGAGTGTGCCATCATCACGACTTCCCTGGACAGCGGCGACGGTGTGCTGCTCTG 1272
Db 1300 CGCTGACGCGCGCATGTATACAGGTTCCTTGTACCGGTGGGGCTGTGCTGTGAC 1359
QY 1273 GACCAACCCAGCAAGCCC---ATCTCCCTGCCCGAGGATCTGCCGGGCGCCAGCTACACC 1329
Db 1360 GACCTCTCTGCCAAGGACATTTATGACTTCCCTCGGTGCCACCTGGCTCTCTATGAT 1419
QY 1330 CTGAGCCAGCAGTGGAGCTGGCTTTTGGGTGGGCTCCAGCCCTGTCTTACATGAG 1389
Db 1420 GTAAGCCACAGTGGCGCTCCAGTACGGGCTTACTCTGCTCTGCGAGGACATGAT 1479
QY 1390 TA---CTGCACCAAGCTGTGTGTACCGCGGAGGCGCAAGGACAGATGTTGCCAGACC 1446
Db 1480 ATGTCTGCCACACTCTGTGTGCTCTGTGGGACCA-----CCTGTGCTACCTCC 1527
QY 1447 CGCCACTTCCCTGGGCGCATGGCACAGCTGTGGCGAGGCAAGCTCTGCTCAAGGG 1506
Db 1528 AAGCTGGATGCAAGCTGTGGAGCGCACCGGTGTGGGGAGATAAGTGTGTCTCAGTGG 1587
QY 1507 GCCTGCTGGAGAGACAACTCAACAAAGCACAGGTTGGATGTTCTTGGGCCAAATGG 1566
Db 1588 GAGTGGTACCGTGGGCTTCCGGCCCGAGGCC---GTGGATGTTGGTGTGCTGTGGCTGG 1644
QY 1567 GATCCCTATGCTGCTCGGCGCATGTGTGGGGCGGTGACAGTGGCCAGAGCGAG 1626
Db 1645 AGCGCTGTCTCATCTGCTACGAGCTGTGGCATGGCGGTACAGAGCGCGAGCGGACG 1704
QY 1627 TGCACAAACCCCATCCCTGCCAAGGGGGCAAGTACTCGAGGAGGTGAGGCTGAATAC 1686
Db 1705 TGCAGCAGCTACGCCCAATACAAAGGAGATGACTGTGTGGGTGAGCGCAAGCGTTC 1764
QY 1687 CGATCTGCAATCTGGAGCCCTGCCAGCTGACGCTCCGGAAGAGCTTCCGGGAGAG 1746
Db 1765 CGCTCTCAACCTGACGGCTGCC-----CTGTGCGCCCGCTCCTCTCCGCGCACTC 1818
QY 1747 CAGTGTGAGGCTTCAACGGGTACACACAGCAGCAACCGGCTCACTCTGCGCGTTCACA 1806
Db 1819 CAGTGCAGCCACTTTGACGCTATGCTCTACAAGGCGGACGTGCACA-----CA 1866
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 16:06:39 ; Search time 576 Seconds
(without alignments)
11154.429 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgctctgctggcgcctc.....gcgtcctgagggcgtgctga 2853

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	2853	100.0	2853 24	AAD35569	Human protease CDN
2	2853	100.0	3446 24	AAD35571	Human protease CDN
3	2846.6	99.8	2853 22	AAH41003	Human metalloprote
4	2817.6	98.8	2930 24	ABK12894	Human protease PRT
5	2817.6	98.8	2937 24	AAH97182	Human metalloprote
6	1215	42.6	1518 21	AAA95827	Rat metalloprote
7	1091.2	38.2	1104 24	AAD35570	Human protease CDN
8	1043.6	36.6	1143 21	AAA95831	Human metalloprote
9	959	33.6	966 24	AAD35568	Human protease CDN

10	664.4	23.3	2670	22	AAH20226	Human ADAM-type me
11	664.4	23.3	2670	22	AAH20216	Human ADAM-type me
12	664.4	23.3	3008	20	AAH20201	Human METH2 encodi
13	664.4	23.3	3008	22	AAH20201	Human METH2 coding
14	634.4	22.2	4180	20	AAH20210	Human METH1 relate
15	634.4	22.2	4180	22	AAH20210	D67076 cDNA clone.
16	628.2	22.0	3261	22	AAH20200	Human METH1 encodi
17	628.2	22.0	3261	22	AAH20200	Human METH1 coding
18	628.2	22.0	4014	19	AAH20200	Human integrin lig
19	627.2	22.0	2184	20	AAH17990	Human ADAMTS-1 cod
20	627.2	22.0	4712	20	AAH17990	Human ADAMTS-1 cod
21	626.6	22.0	2853	22	AAH20224	Human secreted pro
22	626.6	22.0	4676	20	AAH19955	Human ADAM-type me
23	617.2	21.6	4858	24	ABH86234	Human Tango-71 enc
24	607.2	21.3	2346	21	AAA95821	Human metalloprote
25	598.2	21.0	3126	24	ABH86300	Human metalloprote
26	590.4	20.7	3638	22	AAH20224	Human metalloprote
27	584.8	20.5	4192	20	AAH20224	Murine ADAMTS-8 CD
28	583.2	20.4	4301	22	AAH20224	Human aggrecanase-
29	583.2	20.4	4303	21	AAH20224	Human aggrecanase-
30	581.6	20.4	4406	22	AAH20224	Human metalloprote
31	581.6	20.4	4407	21	AAH20224	Human metalloprote
32	581.6	20.4	4407	22	AAH20224	Human metalloprote
33	549	19.2	3250	20	AAH20224	Probe #46 used in
34	547.4	19.2	3002	22	AAH20224	Murine aggrecan deg
35	547.4	19.2	5530	21	AAH20224	Murine ADAMTS-5 CD
36	528.6	18.5	1878	24	ABH86276	Human aggrecanase co
37	524.6	18.4	2751	21	AAH20224	Rat metalloprote
38	469.8	16.5	2625	22	AAH20224	Murine ADAMTS-9 CD
39	419.2	14.7	5053	24	ABH88916	Human aggrecanase
40	417.6	14.6	5605	24	ABH88915	Human aggrecanase
41	417.6	14.6	5808	24	AAH97179	Human metalloprote
42	417.6	14.6	5949	24	AAH98771	Human ADAMTS-SI nu
43	417.6	14.6	6033	24	AAH20224	Human metalloprote
44	416	14.6	4086	22	AAH20224	Human metalloprote
45	409.6	14.4	5804	22	AAH20224	Human ADAMTS-9 CD

ALIGNMENTS

RESULT 1

AAD35569
ID AAD35569 standard; cDNA; 2853 BP.

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX AC AAD35569;

XX WPI; 2002-372123/40.
DR P-PSDB; AAD22541.
XX
PT Novel nucleic acid encoding a human protease, useful as a hybridization
PT probe for screening libraries and assessing gene expression patterns -
XX
PS Claim 1; Page 35-36; 4lpp; English.
XX
CC The present sequence is a cDNA encoding novel human protein (NHP),
CC human protease. NHPs share structural similarity with animal proteases
CC particularly zinc metalloproteases. Sequences of the invention are
CC useful in therapeutic, diagnostic and pharmacogenomic applications.
CC NHP polynucleotides are used as hybridisation probes for screening
CC libraries and assessing gene expression patterns. They can also be
CC used for treating related biological disorders such as obesity, high
CC blood pressure, arthritis, connective tissue disorders and infertility.
CC They are also used in gene therapy.
XX
SQ Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other;
Query Match 100.0%; Score 2853; DB 24; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTTCTGCTGGGCAATCAACCTGGCTTCGCGGGCGAACCCTGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGCAATCAACCTGGCTTCGCGGGCGAACCCTGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTGCTCCATCCGACTGGACCGCGACATTAACGCGCGCGCTAC 120
DB 61 CCAGAGCGGAGGTAGTCTGCTCCATCCGACTGGACCGCGACATTAACGCGCGCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCAGACATTT 180
DB 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCAGACATTT 180
QY 181 CAGGAGGACTTTTACCTACACCTGAGCGCGGATGCTCAGTTTGGCTCCGCGCTTCTCC 240
DB 181 CAGGAGGACTTTTACCTACACCTGAGCGCGGATGCTCAGTTTGGCTCCGCGCTTCTCC 240
QY 241 ACTGAGCATCTGGCGTCCCTCCAGGGCTCACCGGGGCTTTCAGACCTGGCGACGC 300
DB 241 ACTGAGCATCTGGCGTCCCTCCAGGGCTCACCGGGGCTTTCAGACCTGGCGACGC 300
QY 301 TGTCTTATTCTGGGACGTAAACCGCGAGCGAGCTCGTTCGCTGCTGTGAGCCTGTGC 360
DB 301 TGTCTTATTCTGGGACGTAAACCGCGAGCGAGCTCGTTCGCTGCTGTGAGCCTGTGC 360
QY 361 GGGGGCTCCGCGAGCCTTTGGTACCGAGCGCGAGTATGTATAGCCCGCTGCC 420
DB 361 GGGGGCTCCGCGAGCCTTTGGTACCGAGCGCGAGTATGTATAGCCCGCTGCC 420
QY 421 AATCTAGCGCGCGCGCGAGCGCAACAGCGCGACACCTTTCAGCGCGCGG 480
DB 421 AATCTAGCGCGCGCGCGAGCGCAACAGCGCGAGCGCAACCTTTCAGCGCGCGG 480
QY 481 GGTGTTCCGGCGGGCTTCCGAGAGCCCACTCTGCTGCGGGGTGGCTTCGGGCTG 540
DB 481 GGTGTTCCGGCGGGCTTCCGAGAGCCCACTCTGCTGCGGGGTGGCTTCGGGCTG 540
QY 541 AACCCCGCATCTACGGGCTTGAGCCCTTACAAGCGCGGGGGGCTTCGGGGAG 600
DB 541 AACCCCGCATCTACGGGCTTGAGCCCTTACAAGCGCGGGGGGCTTCGGGGAG 600
QY 601 AGTCTAGCGCGCGAGGCTTGGCGGGCCAAAGGTTTCTGTCTATCCCGCGGTACGTG 660
DB 601 AGTCTAGCGCGCGAGGCTTGGCGGGCCAAAGGTTTCTGTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGTGTGCGCGAGGAGTCAATGTCAAGTTCACGCGCGGACCTTGGAAACAT 720
DB 661 GAGACGCTGTGTGCGCGAGGAGTCAATGTCAAGTTCACGCGCGGACCTTGGAAACAT 720

QY 721 TATCTGTGAGCTGCTGGCAACGGCGGCGGACTTACCGCCATCCAGCATCTCTCAAC 780
DB 721 TATCTGTGAGCTGCTGGCAACGGCGGCGGACTTACCGCCATCCAGCATCTCTCAAC 780
QY 781 CCCATCAACATCGTGTGGTCAAGGTGCTCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
DB 781 CCCATCAACATCGTGTGGTCAAGGTGCTCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
QY 841 GTCACCGCAATCGGGCCCTGACGCTGGCAACTTCTGTGCTGGCAGAGAGAGCTGAAC 900
DB 841 GTCACCGCAATCGGGCCCTGACGCTGGCAACTTCTGTGCTGGCAGAGAGAGCTGAAC 900
QY 901 AAATGAGTACAGACACCCGAGTACTGGGACACTGCCATCTCTTCCACGAGGAGGAC 960
DB 901 AAATGAGTACAGACACCCGAGTACTGGGACACTGCCATCTCTTCCACGAGGAGGAC 960
QY 961 CTGTGTGAGGACCAACCTGTGACACCTGGGATGGCTGATGTGGGTACCATGTGTGAC 1020
DB 961 CTGTGTGAGGACCAACCTGTGACACCTGGGATGGCTGATGTGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGCTGCTGTGCTATTGAGGACATGGGCTTCCATCAGGCTTCCACACTGCC 1080
DB 1021 CCCAAGAGAGCTGCTGTGCTATTGAGGACATGGGCTTCCATCAGGCTTCCACACTGCC 1080
QY 1081 CACGAGCTGGCGGACGCTGTTCAACATGCCCATGCCATGCAATGTGAAAGTCTGTGAGAGGTG 1140
DB 1081 CACGAGCTGGCGGACGCTGTTCAACATGCCCATGCCATGCAATGTGAAAGTCTGTGAGAGGTG 1140
QY 1141 TTTGGAAAGCTCCGAGCAACACATGATGTCCCGGACCTTCATCCAGATCGACCGTGC 1200
DB 1141 TTTGGAAAGCTCCGAGCAACACATGATGTCCCGGACCTTCATCCAGATCGACCGTGC 1200
QY 1201 AACCCCTGGTCAAGTGCAGTGCCTATCATCAGGCTTCCCTGGACAGCGGACGGT 1260
DB 1201 AACCCCTGGTCAAGTGCAGTGCCTATCATCAGGCTTCCCTGGACAGCGGACGGT 1260
QY 1261 GACTGCTTCTGGAGCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTCCCGGGCGCC 1320
DB 1261 GACTGCTTCTGGAGCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTCCCGGGCGCC 1320
QY 1321 AGCTACACCTGAGCAGCAGTGGCTGCTTTTGGCGTGGGCTCCAAAGCCTGTGCT 1380
DB 1321 AGCTACACCTGAGCAGCAGTGGCTGCTTTTGGCGTGGGCTCCAAAGCCTGTGCT 1380
QY 1381 TACATGAGTACTGACCAAGCTGTGTGTCACCGGAGGCAAGGACACAGTGGTGC 1440
DB 1381 TACATGAGTACTGACCAAGCTGTGTGTCACCGGAGGCAAGGACACAGTGGTGC 1440
QY 1441 CAGACCCGCTTCTCCCTGGGCGCATGGCAGCAGCTGTGGCGAGGCAAGCTCTGCTC 1500
DB 1441 CAGACCCGCTTCTCCCTGGGCGCATGGCAGCAGCTGTGGCGAGGCAAGCTCTGCTC 1500
QY 1501 AAAGGGCTGCTGGGAGAGACAACTCAACAGCAGAGGCTGGATGTCTCTGGGCT 1560
DB 1501 AAAGGGCTGCTGGGAGAGACAACTCAACAGCAGAGGCTGGATGTCTCTGGGCT 1560
QY 1561 AAATGGATTCCTATGGCCCTGCTCGGCACATGTGTGGGGGCTGCAGCTGGCCAGG 1620
DB 1561 AAATGGATTCCTATGGCCCTGCTCGGCACATGTGTGGGGGCTGCAGCTGGCCAGG 1620
QY 1621 AGGCAGTGCACCAACCCACCTGCCAAGCGGGCAAGTACTGCGAGGAGTGGGGTG 1680
DB 1621 AGGCAGTGCACCAACCCACCTGCCAAGCGGGCAAGTACTGCGAGGAGTGGGGTG 1680
QY 1681 AAATACCGATTCGCAATCTGAGGCTTCCAGCGCTACACACACCAACCGGCTCCTCTGCC 1740
DB 1681 AAATACCGATTCGCAATCTGAGGCTTCCAGCGCTACACACACCAACCGGCTCCTCTGCC 1740
QY 1741 GAGGAGCAGTGTGAGGCTTCAACGGCTACACACACCAACCGGCTCCTCTGCC 1800
DB 1741 GAGGAGCAGTGTGAGGCTTCAACGGCTACACACACCAACCGGCTCCTCTGCC 1800
QY 1801 GTGGCATGGTGGCCCAAGTACTTCCCGCGGTGTCTCCCGGGAGCAAGTCAAGCTCATCTGC 1860

```
Db 1801 GTGGCTGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGCAAGTCAAGCTCATCTGCG 1860
QY 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGGTGGACGGCAAGCTG 1920
Db 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGGTGGACGGCAAGCTG 1920
QY 1921 TGCTCTCTGACTTCCCACTCCGCTCTGTGTCCAAAGCAAGTGCATCAAGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTTCCCACTCCGCTCTGTGTCCAAAGCAAGTGCATCAAGCTGGCTGTGAT 1980
QY 1981 GGGAACTGGCTTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGCTTCCAAAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGTGAAGTGAAGTGTCCCAAGCAAGTGCATCAAGCTGGCTGTGATG 2100
Db 2041 AGCTGCAAGAAGTGAAGTGAAGTGTCCCAAGCAAGTGCATCAAGCTGGCTGTGATG 2100
QY 2101 GCCATCCCGCAGGGCCCTCAAGCATCGACATCCCGCAGCGGTTACAAAGGGCTGATC 2160
Db 2101 GCCATCCCGCAGGGCCCTCAAGCATCGACATCCCGCAGCGGTTACAAAGGGCTGATC 2160
QY 2161 GGGGATGACAATCTACTGGCTCTGAAGACAGCAAGGCAAGTACCTGCTCAACGGGCAAT 2220
Db 2161 GGGGATGACAATCTACTGGCTCTGAAGACAGCAAGGCAAGTACCTGCTCAACGGGCAAT 2220
QY 2221 TTCTGGTGTCTGGCGGTGGAGCGGACCTGTGGTGTGAAGTCTCTCTCTCTCTCTCTCT 2280
Db 2221 TTCTGGTGTCTGGCGGTGGAGCGGACCTGTGGTGTGAAGTCTCTCTCTCTCTCTCTCT 2280
QY 2281 GGCACGGGCACACGCGGTGGAGAGCTGTGAGGCTTCCCGGCGCATCTCTGAGCGGTGACC 2340
Db 2281 GGCACGGGCACACGCGGTGGAGAGCTGTGAGGCTTCCCGGCGCATCTCTGAGCGGTGACC 2340
QY 2341 GTGGAGTCTCTCTCTGGGGAGATGACACCGCCCGGGTCTCTCTCTCTCTCTCTCTCT 2400
Db 2341 GTGGAGTCTCTCTCTGGGGAGATGACACCGCCCGGGTCTCTCTCTCTCTCTCTCTCT 2400
QY 2401 CCCAAGAGCTCTGGGAGACAAAGTCTCTATCCCAAGAGACCCCGGGGACCTCTCTCT 2460
Db 2401 CCCAAGAGCTCTGGGAGACAAAGTCTCTATCCCAAGAGACCCCGGGGACCTCTCTCT 2460
QY 2461 TTGCACACAGGCTCTCAGCTCTTCCAAACAGGTGGAGTGGAGCGGAGCAGACGCGCCCT 2520
Db 2461 TTGCACACAGGCTCTCAGCTCTTCCAAACAGGTGGAGTGGAGCGGAGCAGACGCGCCCT 2520
QY 2521 GCACGCTGGTGGCTGGAGCTGGGGCGGCTCTCCCGGAGCTGGGGCGAGTGGCGCTGAG 2580
Db 2521 GCACGCTGGTGGCTGGAGCTGGGGCGGCTCTCCCGGAGCTGGGGCGAGTGGCGCTGAG 2580
QY 2581 AAGCGGCGGTGGAGTGTGGGGGTCTCCCGGGGAGCGGAGCTCCCTGCTGTGTGATGCA 2640
Db 2581 AAGCGGCGGTGGAGTGTGGGGGTCTCCCGGGGAGCGGAGCTCCCTGCTGTGTGATGCA 2640
QY 2641 GCCATCGGCGGTGGAGACACAGCTGTGGGGGAGCGCCGCCCGACCTGGGAGCTCAGC 2700
Db 2641 GCCATCGGCGGTGGAGACACAGCTGTGGGGGAGCGCCGCCCGACCTGGGAGCTCAGC 2700
QY 2701 GCCTGCTCACCTCTCTCAAGAGCTGGCGCGGGGATTTAGAGCGCTCTACTCAAGTGT 2760
Db 2701 GCCTGCTCACCTCTCTCAAGAGCTGGCGCGGGGATTTAGAGCGCTCTACTCAAGTGT 2760
QY 2761 GTGGGCCACGAGGCGGCTGCTGGCCCGGAGCCAGTGCACCTGCACTGCAAGCCGAGCC 2820
Db 2761 GTGGGCCACGAGGCGGCTGCTGGCCCGGAGCCAGTGCACCTGCACTGCAAGCCGAGCC 2820
QY 2821 GAGCTGGACTTCTGCTGCTGAGCGCGTGTGA 2853
Db 2821 GAGCTGGACTTCTGCTGCTGAGCGCGTGTGA 2853
```

RESULT 2

AAD35571

ID AAD35571 standard; cDNA; 3446 BP.

XX

AC AAD35571;

XX

XX 26-JUL-2002 (first entry)

DT

XX

XX Human protease cDNA #4.

DE

XX

XX Human; novel human protein; NHP; protease; biological disorder; obesity;

KW high blood pressure; arthritis; connective tissue disorder; infertility;

KW gene therapy; enzyme; ss.

XX

OS Homo sapiens.

XX

XX WO200226949-A2.

PN

XX 04-APR-2002.

XX

XX 27-SEP-2001; 2001WO-US30350.

XX

XX 29-SEP-2000; 2000US-236689P.

PR

XX (LEXI-) LEXICON GENETICS INC.

XX

XX Friddle CJ, Hilbun E;

PI

XX WPI; 2002-372123/40.

XX

XX Novel nucleic acid encoding a human protease, useful as a hybridization

PT probe for screening libraries and assessing gene expression patterns -

XX

PS Disclosure; Page 40-41; 41pp; English.

XX

CC The present sequence is a cDNA encoding novel human protein (NHP),

CC human protease. NHPs share structural similarity with animal proteases

CC particularly zinc metalloproteases. Sequences of the invention are

CC useful in therapeutic, diagnostic and pharmacogenomic applications.

CC NHP polynucleotides are used as hybridisation probes for screening

CC libraries and assessing gene expression patterns. They can also be

CC used for treating related biological disorders such as obesity, high

CC blood pressure, arthritis, connective tissue disorders and infertility.

CC They are also used in gene therapy.

XX

SQ Sequence 3446 BP; 612 A; 1114 C; 1101 G; 619 T; 0 other;

Query Match 100.0%; Score 2853; DB 24; Length 3446;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGCATCCCTAACCCCTGGCTTTTCGCGGGGGAACCCCTGGAGGCTCTGAG 60

Db 397 ATGCTTCTGCTGGCATCCCTAACCCCTGGCTTTTCGCGGGGGAACCCCTGGAGGCTCTGAG 456

QY 61 CCAGAGCGGAGGTAGTTCGTTCCCATCCGACTGGACCCGCGGACATTAAACGCGCCGCTAC 120

Db 457 CCAGAGCGGAGGTAGTTCGTTCCCATCCGACTGGACCCGCGGACATTAAACGCGCCGCTAC 516

QY 121 TACTGGGGGTCGCGGAGGACTCCGGGAGTCAATTTTTCAGATCACAGCATTT 180

Db 517 TACTGGGGGTCGCGGAGGACTCCGGGAGTCAATTTTTCAGATCACAGCATTT 576

QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 240

Db 577 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGGCTCCGCGCTTCTCC 636

QY 241 ACTGAGCATCTGGCGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACCTGGGAGCG 300

Db 637 ACTGAGCATCTGGCGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACCTGGGAGCG 696

QY 301 TGCTTCTATTCTGGGGACGTGAACCGCGGAGTCTGCTGCTGAGCGCTGTCG 360

Db 697 TGCTTCTATTCTGGGGACGTGAACCGCGGAGTCTGCTGCTGAGCGCTGTCG 756

QY 361 GGGGGCTCCGGAGCCCTTTGGCTACCGAGCGCGAGTATGTCTATAGCCCGCTGCC 420
Db 757 GGGGGCTCCGGAGCCCTTTGGCTACCGAGCGCGAGTATGTCTATAGCCCGCTGCC 816
QY 421 AATGCTAGCGCGCGCGCGAGCGCAAGCAAGCAAGCGCGCGAGTATGTCTATAGCCCGCG 480
Db 817 AATGCTAGCGCGCGCGCGAGCGCAAGCAAGCAAGCGCGCGAGTATGTCTATAGCCCGCG 876
QY 481 GGTGTTCCGGGGGGGGCTTTCCGGAGACCCACCTCTCGCTGCGGGTGGCTTCGGGGTGG 540
Db 877 GGTGTTCCGGGGGGGGCTTTCCGGAGACCCACCTCTCGCTGCGGGTGGCTTCGGGGTGG 936
QY 541 AACCCCGCATCTACGGCCCTTGGACCTTACAGCCGCGCGCGCGCGAGTATGTCTATAGCCCGCG 600
Db 937 AACCCCGCATCTACGGCCCTTGGACCTTACAGCCGCGCGCGCGAGTATGTCTATAGCCCGCG 996
QY 601 AGTCGTAGCGCGCGCGAGTCTGGCGCGCGCAAGCGCTTCTATAGCCCGCGCGAGTATGTCT 660
Db 997 AGTCGTAGCGCGCGCGAGTCTGGCGCGCGCAAGCGCTTCTATAGCCCGCGCGAGTATGTCT 1056
QY 661 GAGACGCTGGTGTGCGGACGAGTCAATGGTCAAGTTCACCGCGCGAGCTTGAACAT 720
Db 1057 GAGACGCTGGTGTGCGGACGAGTCAATGGTCAAGTTCACCGCGCGAGCTTGAACAT 1116
QY 721 TATCTGTACGCTGTGCAAGCGCGCGAGTCTACCGCGCATCCAGCATCTCTCAAC 780
Db 1117 TATCTGTACGCTGTGCAAGCGCGCGAGTCTACCGCGCATCCAGCATCTCTCAAC 1176
QY 781 CCCATCAACATCTTGTGGTCAAGTGTCTTCTTAGAGATCGTACTCCGGCGCGCAAG 840
Db 1177 CCCATCAACATCTTGTGGTCAAGTGTCTTCTTAGAGATCGTACTCCGGCGCGCAAG 1236
QY 841 GTACCGGCAATCGGCCCTGACGCTGCGCAACTTCTGTGCGTGGCAGAAAGTGAAC 900
Db 1237 GTACCGGCAATCGGCCCTGACGCTGCGCAACTTCTGTGCGTGGCAGAAAGTGAAC 1296
QY 901 AAAGTGTAGTACAAAGCAACCCCGAGTACTGGGACACTGCCATCTCTTACAGGCGAGGAC 960
Db 1297 AAAGTGTAGTACAAAGCAACCCCGAGTACTGGGACACTGCCATCTCTTACAGGCGAGGAC 1356
QY 961 CTGTGTGGAGCCACCACTGTGACACCTGGGATGGCTATGTGGTACCTATGTGTGAC 1020
Db 1357 CTGTGTGGAGCCACCACTGTGACACCTGGGATGGCTATGTGGTACCTATGTGTGAC 1416
QY 1021 CCCAAGAAAGTGTCTGTCTATTGAGGAGTGGCTTCCATCAGCCTTCCACACTGCC 1080
Db 1417 CCCAAGAAAGTGTCTGTCTATTGAGGAGTGGCTTCCATCAGCCTTCCACACTGCC 1476
QY 1081 CACGAGCTGGGCGACGTTTCAACATGCCCCCATGACAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1477 CACGAGCTGGGCGACGTTTCAACATGCCCCCATGACAATGTGAAGTCTGTGAGGAGGTG 1536
QY 1141 TTTGGGAAGCTCCGAGCCACCACTGATGTCCCGACCTCATCCAGATCGACCGTGC 1200
Db 1537 TTTGGGAAGCTCCGAGCCACCACTGATGTCCCGACCTCATCCAGATCGACCGTGC 1596
QY 1201 AACCCCTGGTCAAGCTGTGCTGATCATCAGGACTTCTTGGACAGCGGCGACGCT 1260
Db 1597 AACCCCTGGTCAAGCTGTGCTGATCATCAGGACTTCTTGGACAGCGGCGACGCT 1656
QY 1261 GACTGCTCTGTGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGCCGGCGCC 1320
Db 1657 GACTGCTCTGTGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGCCGGCGCC 1716
QY 1321 AGCTACACCTGTAGCCAGAGTGGAGTGGCTTTTGGCTGGGCTTCCAGGCGCTGTCT 1380
Db 1717 AGCTACACCTGTAGCCAGAGTGGAGTGGCTTTTGGCTGGGCTTCCAGGCGCTGTCT 1776
QY 1381 TACATGAGTACTGACCAAGAGTGTGTGCAACCGGGAAGGCCAAGGAGACAGATGGTGTGC 1440
Db 1777 TACATGAGTACTGACCAAGAGTGTGTGCAACCGGGAAGGCCAAGGAGACAGATGGTGTGC 1836

QY 1441 CAGACCGCCACCTTCCCTGGGCGGATGGACACGAGTGTGGCGAGGCAAGCTCTGCCIC 1500
Db 1837 CAGACCGCCACCTTCCCTGGGCGGATGGACACGAGTGTGGCGAGGCAAGCTCTGCCIC 1896
QY 1501 AAAGGGCCCTCGGTGGAGAGACACAACCTCAACAAGCACAGGTGGATGTTCTTCTGGGC 1560
Db 1897 AAAGGGCCCTCGGTGGAGAGACACAACCTCAACAAGCACAGGTGGATGTTCTTCTGGGC 1956
QY 1561 AAATGGGATCCCTATGTCCTGCTCGCGACATGTGGTGGGGCGGTGCAAGCTGGCAAG 1620
Db 1957 AAATGGGATCCCTATGTCCTGCTCGCGACATGTGGTGGGGCGGTGCAAGCTGGCAAG 2016
QY 1621 AGGCAGTGCACCAACCCACCTGTCACAGGGGCAAGTACTGCGAGGAGTGAAGGIG 1680
Db 2017 AGGCAGTGCACCAACCCACCTGTCACAGGGGCAAGTACTGCGAGGAGTGAAGGIG 2076
QY 1681 AAATACCGATCTCAATCTGGAGCCTGCTCCAGCTCAGCTCCCGAAAGCTTCCGG 1740
Db 2077 AAATACCGATCTCAATCTGGAGCCTGCTCCAGCTCAGCTCCCGAAAGCTTCCGG 2136
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACGGCTTCAACGGCTTCACTCTCGCC 1800
Db 2137 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACGGCTTCAACGGCTTCACTCTCGCC 2196
QY 1801 GTGGATGGGTGCCCAAGTACTCGGGGTGCTCCCCGGGACAGTGCACAGCTCATCTGC 1860
Db 2197 GTGGATGGGTGCCCAAGTACTCGGGGTGCTCCCCGGGACAGTGCACAGCTCATCTGC 2256
QY 1861 CGAGCAATGCACTGGCTACTTCTATGTGTGCAACCCCAAGTGTGGAGCGGACGCTG 1920
Db 2257 CGAGCAATGCACTGGCTACTTCTATGTGTGCAACCCCAAGTGTGGAGCGGACGCTG 2316
QY 1921 TGCTCTCTGACTCCACCTCGCTGTGTCACAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 2317 TGCTCTCTGACTCCACCTCGCTGTGTCACAGGCAAGTGCATCAAGGCTGGCTGTGAT 2376
QY 1981 GGGACCTGGCTCCACAGAGATTCGCAAGTGTGGGTGTGTGGGGAGACAATAAG 2040
Db 2377 GGGACCTGGCTCCACAGAGATTCGCAAGTGTGGGTGTGTGGGGAGACAATAAG 2436
QY 2041 AGCTGCAAGAAAGTGTGACTCTTCAACAGCCCATGCAATGGCTACAAATTCGTGGTG 2100
Db 2437 AGCTGCAAGAAAGTGTGACTCTTCAACAGCCCATGCAATGGCTACAAATTCGTGGTG 2496
QY 2101 GCATPCCCGCAGGGCCCTCAAGCATCGCATCCGCGAGCGGTTTCAAGAGGCTCATC 2160
Db 2497 GCATPCCCGCAGGGCCCTCAAGCATCGCATCCGCGAGCGGTTTCAAGAGGCTCATC 2556
QY 2161 GGGGATGCAACTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGGCAT 2220
Db 2557 GGGGATGCAACTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGGCAT 2616
QY 2221 TTCTGTGTGTGGGGTGGAGCGGACCTGTGTGTAAGGCGAGTGTGCTGCGGTACAGC 2280
Db 2617 TTCTGTGTGTGGGGTGGAGCGGACCTGTGTGTAAGGCGAGTGTGCTGCGGTACAGC 2676
QY 2281 GGCACGGGACAGCGGTGGAGCGCTGAGGCTTCCCGGCCCATCTTGGAGCGGCTGACC 2340
Db 2677 GGCACGGGACAGCGGTGGAGCGCTGAGGCTTCCCGGCCCATCTTGGAGCGGCTGACC 2736
QY 2341 GTGGAGGTCTCTCGGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2737 GTGGAGGTCTCTCGGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2796
QY 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTATCCAGGACCCCGGGGACCTCTGTGC 2460
Db 2797 CCCAAGAGCCTCGGAGGACAAAGTCTCTATCCAGGACCCCGGGGACCTCTGTGC 2856
QY 2461 TTGCAACACAGCGCTCTCAGCTCTCCAGCAGGTGGAGCGGACGAGCAGGCGCCCT 2520
Db 2857 TTGCAACACAGCGCTCTCAGCTCTCCAGCAGGTGGAGCGGACGAGCAGGCGCCCT 2916
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGGCGGCTCTCCGAGGCTGGCGAGTGGGCTGCAG 2580

```

Db 2917 GCACGCTGGTGGCTGGCAGCTGGGGCGGTGCTCCGGGAGCTGGGAGTGGCTGCAG 2976
QY 2581 AAGCGGGGGTGGAGTGTCTGGGGCTCCCGGGGAGCGCACGGTCCCTGTGTATGCA 2640
Db 2977 AAGCGGGGGTGGAGTGTCTGGGGCTCCCGGGGAGCGCACGGTCCCTGTGTATGCA 3036
QY 2641 GCCCATCGCCCGTGGAGACACAAGCCCTGCGGGGAGCCCTGCCACCTGGAGCTCAGC 2700
Db 3037 GCCCATCGCCCGTGGAGACACAAGCCCTGCGGGGAGCCCTGCCACCTGGAGCTCAGC 3096
QY 2701 GCCTGGTCAACCTGTCTCAAGAGTGTGCGCCGGGGATTTACAGCGGCTCACTCAAGTGT 2760
Db 3097 GCCTGGTCAACCTGTCTCAAGAGTGTGCGCCGGGGATTTACAGCGGCTCACTCAAGTGT 3156
QY 2761 GTGGCCACGAGCGCGGCTGCTGGCCGGGACCAAGTGTCACTTGCACCCGACCCAG 2820
Db 3157 GTGGCCACGAGCGCGGCTGCTGGCCGGGACCAAGTGTCACTTGCACCCGACCCAG 3216
QY 2821 GAGCTGGACTTCTGCTGCTCTGAGCGCGTGTGA 2853
Db 3217 GAGCTGGACTTCTGCTGCTCTGAGCGCGTGTGA 3249

RESULT 3
AAH41003
ID AAH41003 standard; cDNA; 2853 BP.
XX AC AAH41003;
XX DT 23-AUG-2001 (first entry)
XX DE Human metalloprotease MDT56 cDNA.
XX KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
XX KW osteopathic; antiarthritic; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 1..2853
XX FT /*tag= a
XX FT /product= "Metalloprotease"
XX FT /note= "Aggrecanase activity"
XX PN W0200134785-A1.
XX PD 17-MAY-2001.
XX PF 10-NOV-2000; 2000WO-JP07917.
XX PR 11-NOV-1999; 99JP-0321740.
XX PR 16-MAY-2000; 2000JP-0144020.
XX PA (YAMA ) YAMANOUCHI PHARM CO LTD.
XX PA (KAZU-) KAZUSA DNA RES INST.
XX PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
XX WPI; 2001-343602/36.
XX DR P-PSDB; AAG62299.
XX PT Metalloprotease with aggrecanase activity for treating joint diseases
XX PT especially osteoarthritis -
XX PS Example 2; Page 60-61; 85pp; Japanese.
XX CC This invention relates to a metalloprotease with aggrecanase activity.
XX CC The invention includes protein and DNA sequences of the metalloprotease,
XX CC vectors containing the DNA, host cells transformed by the vectors, and
XX CC antibodies directed against the metalloprotease. The antibodies, protein
XX CC and DNA sequences can be used in the treatment and prevention of joint
XX CC diseases, particularly osteoarthritis. The treatment may result in

```

CC osteopathic and antiarthritic activity. The present sequence represents
 CC cDNA encoding the metalloprotease termed MDT56.

XX
 SQ Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other;

Query Match 99.8%; Score 2846.6; DB 22; Length 2853;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2849; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATGCTTCTGCTGGCATCTTAACCTTGGCTTTCGCCGGGGAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGCATCTTAACCTTGGCTTTCGCCGGGGAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGTAGTTCGTTCCATCCGACTGGACCCGCGACATTAAACGCGCGCGTAC 120
Db 61 CCAGAGCGGAGTAGTTCGTTCCATCCGACTGGACCCGCGACATTAAACGCGCGCGTAC 120
QY 121 TACTGGGGGTCGCCAGGACTCCGGGGATCAGGACTCATTTTTCAGATCACAGCAIT 180
Db 121 TACTGGGGGTCGCCAGGACTCCGGGGATCAGGACTCATTTTTCAGATCACAGCAIT 180
QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCCGCTTC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCCGCTTC 240
QY 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTCACCGGGGGCTCTTTCAGACCTCGCACGC 300
Db 241 ACTGAGCATCTGGGCTCCCTCCAGGGGCTCACCGGGGGCTCTTTCAGACCTCGCACGC 300
QY 301 TGCTTCTATTCTGGGACGTGAACGCCGAGCGGACTCCTTCGCTGCTGTGAGCTGTGC 360
Db 301 TGCTTCTATTCTGGGACGTGAACGCCGAGCGGACTCCTTCGCTGCTGTGAGCTGTGC 360
QY 361 GGGGGCTCCGGGAGCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCGCTGCC 420
Db 361 GGGGGCTCCGGGAGCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCGCTGCC 420
QY 421 AATGCTAGCGCGCGCGCGCACACAGCCAGGGCGCACACTTCTCCAGCGCGCGG 480
Db 421 AATGCTAGCGCGCGCGCGCACACAGCCAGGGCGCACACTTCTCCAGCGCGCGG 480
QY 481 GGTGTTCCGGGGGGCTTCGGAGACCCACCTCTCGTGGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCCGGGGGGCTTCGGAGACCCACCTCTCGTGGGGGTGGCTCGGGCTGG 540
QY 541 AACCCCGCATCTACGGGCTTGGACCTTACAGCGCGCGGCGGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTACGGGCTTGGACCTTACAGCGCGCGGCGGGCTTCGGGGAG 600
QY 601 AGTCTAGCGCGCGAGTCTTGGGCGCGCAAGCGTTTGTGTCTATCCCGGCTACG 660
Db 601 AGTCTAGCGCGCGAGTCTTGGGCGCGCAAGCGTTTGTGTCTATCCCGGCTACG 660
QY 661 GAGACGCTGGTGTGCGGACGAGTCAATGTCTCCAGGGCGGAGACTGGAACAT 720
Db 661 GAGACGCTGGTGTGCGGACGAGTCAATGTCTCCAGGGCGGAGACTGGAACAT 720
QY 721 TATCTGTGACGCTGTGSCAACGCGGCGGACTTACCGCCATCCAGCATCTCAAC 780
Db 721 TATCTGTGACGCTGTGSCAACGCGGCGGACTTACCGCCATCCAGCATCTCAAC 780
QY 781 CCATCAACATGTTGTGTTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCGCCAG 840
Db 781 CCATCAACATGTTGTGTTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCGCCAG 840
QY 841 GTCACGGGAATGCGGCGCTGACGCTGCGCACTTCTGTGCTGGCAGAGAAGCTGAAC 900
Db 841 GTCACGGGAATGCGGCGCTGACGCTGCGCACTTCTGTGCTGGCAGAGAAGCTGAAC 900
QY 901 AAAGTGTGACAGCAACCCCGAGTCTGGGACACTGCGCATCTTCCAGCAGCAGAC 960
Db 901 AAAGTGTGACAGCAACCCCGAGTCTGGGACACTGCGCATCTTCCAGCAGCAGAC 960

```

```
QY 961 CTGTGTGGAGCCACCACTGTGTACACCCCTGGGCATGGCTGTATGTGGGTACCATGTGTGAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGTACACCCCTGGGCATGGCTGTATGTGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGCTGCTGTCTCATTTGAGGAGGATGGGCTTCCATCAGCCCTTACCACCTGCC 1080
Db 1021 CCCAAGAGAAGCTGCTGTCTCATTTGAGGAGGATGGGCTTCCATCAGCCCTTACCACCTGCC 1080
QY 1081 CAGCAGCTGGGCGACCTGTTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1081 CAGCAGCTGGGCGACCTGTTTCAACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCAACCAACATGATGTGTCCCGACCCCTCATTCAGATTCGACCGTGC 1200
Db 1141 TTTGGGAAGCTCCGAGCAACCAACATGATGTGTCCCGACCCCTCATTCAGATTCGACCGTGC 1200
QY 1201 AACCCCTGGTCAAGCTGCTGCTGCATCATCACCAGCTTCTGTGACACGGGACAGGT 1260
Db 1201 AACCCCTGGTCAAGCTGCTGCTGCATCATCACCAGCTTCTGTGACACGGGACAGGT 1260
QY 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCCCGAGGATCTGCCGGGCGCC 1320
Db 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCCCGAGGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTTCCAAAGCCCTGTCT 1380
Db 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTTCCAAAGCCCTGTCT 1380
QY 1381 TACATGCACTGTGACCAAGCTGTGGTGACCGGGAAGGC/AAGGACAGATGTTGTG 1440
Db 1381 TACATGCACTGTGACCAAGCTGTGGTGACCGGGAAGGC/AAGGACAGATGTTGTG 1440
QY 1441 CAGACCCGCACTTCCCTGGGCGGATGGCAACAGTGTGGCGAGGGCAAGCTCTGCCCTC 1500
Db 1441 CAGACCCGCACTTCCCTGGGCGGATGGCAACAGTGTGGCGAGGGCAAGCTCTGCCCTC 1500
QY 1501 AAGGGCGCTTGGTGAGAGACACACCTCAACAAGCAGAGGTGGATGTTCTTGGGCC 1560
Db 1501 AAGGGCGCTTGGTGAGAGACACACCTCAACAAGCAGAGGTGGATGTTCTTGGGCC 1560
QY 1561 AAATGGGATCCCTATGCGCCCTGTGTCGCGACATGTGGTGGGGCGGTGCAAGCTGGCCAG 1620
Db 1561 AAATGGGATCCCTATGCGCCCTGTGTCGCGACATGTGGTGGGGCGGTGCAAGCTGGCCAG 1620
QY 1621 AGGCAGTGCACCAACCCCTGCTGCGCGACATGTGGTGGGGCGGTGCAAGCTGGCCAG 1680
Db 1621 AGGCAGTGCACCAACCCCTGCTGCGCGACATGTGGTGGGGCGGTGCAAGCTGGCCAG 1680
QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCCCTCCGAAAGAGCTTCCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCCCTCCGAAAGAGCTTCCGG 1740
QY 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTACACCAACAGCACCACCGGCTTCACTCTCGCC 1800
Db 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTACACCAACAGCACCACCGGCTTCACTCTCGCC 1800
QY 1801 GTGGCATGGTGCCCAAGTACTCGGGGCTGTCTCCCGGGACAAAGTGCACAGCTCACTCTGC 1860
Db 1801 GTGGCATGGTGCCCAAGTACTCGGGGCTGTCTCCCGGGACAAAGTGCACAGCTCACTCTGC 1860
QY 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGTCGACCCCAAGGTGGTGGACGGCAGGCTG 1920
Db 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGTCGACCCCAAGGTGGTGGACGGCAGGCTG 1920
QY 1921 TGCTCTCTGACTCCACTCCCTGCTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACTCCCTGCTGTGTCCAAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
QY 1981 GGGAACTGGCTCCAGAGAGATTCGACAAGTGTGGGTGTGTGGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGCTCCAGAGAGATTCGACAAGTGTGGGTGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGCTAGCTACAAATTCGTGGTG 2100
```

```
Db 2041 AGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGCTAGCTACAAATTCGTGGTG 2100
QY 2101 GCCATCCCGCAGCGCCTCAAGCATCCGCCAGCGCGGTACAAAGGCTGATC 2160
Db 2101 GCCATCCCGCAGCGCCTCAAGCATCCGCCAGCGCGGTACAAAGGCTGATC 2160
QY 2161 GGGGATGACAACACTACCTGGCTCTGAAGAACAGCAAGTACCTGCTCAACGGCAT 2220
Db 2161 GGGGATGACAACACTACCTGGCTCTGAAGAACAGCAAGTACCTGCTCAACGGCAT 2220
QY 2221 TTCGTGGTGTGCGCGGTGGAGCGGACCTGCTGTGAAGGCGAGTCTGCTGCGGTACAGC 2280
Db 2221 TTCGTGGTGTGCGCGGTGGAGCGGACCTGCTGTGAAGGCGAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCACGGGCACAGCGGTGGAGAGCTGTCAGGCTTCCCGGGCCCATCTCTGGAGCCCTGACC 2340
Db 2281 GGCACGGGCACAGCGGTGGAGAGCTGTCAGGCTTCCCGGGCCCATCTCTGGAGCCCTGACC 2340
QY 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2341 GTGGAGTCTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
QY 2401 CCCAAGAGCTCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCCCTCTGTC 2460
Db 2401 CCCAAGAGCTCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCCCTCTGTC 2460
QY 2461 TTGCAACAACAGCGTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGACAGCGCCCT 2520
Db 2461 TTGCAACAACAGCGTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGACAGCGCCCT 2520
QY 2521 GCACCTGGGTGGTGGCAGCTGGGGGCGGTCTCCCGGAGCTGGCGAGTGGGCTTGCAG 2580
Db 2521 GCACCTGGGTGGTGGCAGCTGGGGGCGGTCTCCCGGAGCTGGCGAGTGGGCTTGCAG 2580
QY 2581 AAGCGGGCGGTGGAGTGTGCGGGGTCTCCCGGGGAGCGCAGCGCTGCTGTGATGCA 2640
Db 2581 AAGCGGGCGGTGGAGTGTGCGGGGTCTCCCGGGGAGCGCAGCGCTGCTGTGATGCA 2640
QY 2641 GCCCATCGGCGGTGGAGACACAGCCCTGCGGGGAGCGCTGCCCGACCTGGGAGCTCAGC 2700
Db 2641 GCCCATCGGCGGTGGAGACACAGCCCTGCGGGGAGCGCTGCCCGACCTGGGAGCTCAGC 2700
QY 2701 GCTGTGTACCTGTCTCCAAAGAGCTGCGGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCTGTGTACCTGTCTCCAAAGAGCTGCGGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCCACGGAGCGCGGTGCTGGCCCGGGACCAAGTGCACCTTGCACCGCAAGCCGAG 2820
Db 2761 GTGGGCCACGGAGCGCGGTGCTGGCCCGGGACCAAGTGCACCTTGCACCGCAAGCCGAG 2820
QY 2821 GAGCTGGACTTCTGGCTCTCTGAGGCGGTGCTGA 2853
Db 2821 GAGCTGGACTTCTGGCTCTCTGAGGCGGTGCTGA 2853
```

RESULT 4
ABK12894

ID ABK12894 standard; cDNA; 2930 BP.

XX ABK12894;

XX 09-APR-2002 (first entry)

XX Human protease PRPS-11 cDNA sequence.

XX Human; protease; PRPS; gastrointestinal; Crohn's disease; cancer;
XX cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
XX inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
XX cell proliferative disorder; developmental disorder; epilepsy;
XX Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
XX reproductive disorder; endometriosis; ss.

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 75..2930
 FT CDS
 FT /*tag= a
 FT /partial
 FT /product= "Human protease PRTS-11"
 FT /note= "This sequence lacks a stop codon"
 XX
 PN W0200198468-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 13-JUN-2001; 2001WO-US19178.
 XX
 PR 16-JUN-2000; 2000US-212336P.
 PR 22-JUN-2000; 2000US-213953P.
 PR 29-JUN-2000; 2000US-215396P.
 PR 07-JUL-2000; 2000US-216821P.
 PR 14-JUL-2000; 2000US-218946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;
 PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kallick DA;
 XX
 DR WPI; 2002-090437/12.
 DR P-PSDB; AAU74751.
 XX
 PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful
 PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell
 PT proliferative (e.g. cancer) disorders -
 XX
 PS Claim 5; Page 168; 177pp; English.
 XX
 CC The present invention relates to twenty one new human proteases,
 CC referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and
 CC polypeptides of the invention are useful in the diagnosis, treatment and
 CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
 CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
 CC myocardial infarction, autoimmune/inflammatory e.g. acquired
 CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
 CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker
 CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
 CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and
 CC endometriosis disorders. Numerous other examples of each disorder are
 CC given in the specification. The present nucleic acid sequence encodes
 CC the human protease PRTS-11 protein of the invention.
 XX
 SQ Sequence 2930 BP; 529 A; 958 C; 931 G; 512 T; 0 other;
 XX
 Query Match 98.8%; Score 2817.6; DB 24; Length 2930;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2846; Conservative 0; Mismatches 4; Indels 6; Gaps 2;
 XX
 QY 1 ATGCTTCTGCTGGGCACTTAACCTGGCTTTCCGGGGGGAACCGCTGGAGGCTCTGAG 60
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 75 ATGCTTCTGCTGGGCACTTAACCTGGCTTTCCGGGGGGAACCGCTGGAGGCTCTGAG 134
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CCAGAGGGGAGGTAGTGTTCATCCGACTGGACCCCGGACATTAAACGGCGCGCTAC 120
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 135 CCAGAGGGGAGGTAGTGTTCATCCGACTGGACCCCGGACATTAAACGGCGCGCTAC 194
 QY 121 TACTGGGGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCATTT 180
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 195 TACTGGGGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCATTT 254
 QY 181 CAGGAGGACTTTTACCTACACCTGACCGGATGCTCAGTTCTTGCTCCCGGCTTCTCC 240
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 255 CAGGAGGACTTTTACCTACACTGACGCCGGATGCTCAGTTCTTGGCTCCCGGCTTCTCC 314
 QY 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACGGGGGCTCTTCCAGACCTTCGACGC 300
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 315 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACGGGGGCTCTTCCAGACCTTCGACGC 374
 QY 301 TGCTTCTATTCTGGGAGCTGAAGCGGAGCCGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 375 TGCTTCTATTCTGGGAGCTGAAGCGGAGCCGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 434
 QY 361 GGGGGCTCCGCGAGGCTTTGGCTACCGAGGCGCGAGTATGCTATATAGCCGCTGCC 420
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 435 GGGGGCTCCGCGAGGCTTTGGCTACCGAGGCGCGAGTATGCTATATAGCCGCTGCC 494
 QY 421 AATGCTAGCGCGCGCGCGCGAGCGCAACAGCCAGGGCGCACACCTTCTTCCAGCGCG 480
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 495 AATGCTAGCGCGCGCGCGCGAGCGCAACAGCCAGGGCGCACACCTTCTTCCAGCGCG 554
 QY 481 GGTGTTCCGGCGCGGCTTCCGAGAGCCGACCTCTGCTGCGGGGTGGCTCGGGCTGG 540
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 555 GGTGTTCCGGCGCGGCTTCCGAGAGCCGACCTCTGCTGCGGGGTGGCTCGGGCTGG 614
 QY 541 AACCCGCGCATCTACGGGCTTGGACCTTACAAGCGCGGGCGGGGCTTCGGGGAG 600
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 615 AACCCGCGCATCTACGGGCTTGGACCTTACAAGCGCGGGCGGGGCTTCGGGGAG 674
 QY 601 AGTCGTAGCGCGCGCGAGGCTTGGCGCGCGCAAGCGTTTGGTGTCTATCCCGGGGTAC 660
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 675 AGTCGTAGCGCGCGAGGCTTGGCGCGCGCAAGCGTTTGGTGTCTATCCCGGGTAC 734
 QY 661 GAGACGCTGGTTCGCGGAGGCTCAATGGTCAAGTTCACGGCGCGGACCTTGGAACT 720
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 735 GAGACGCTGGTTCGCGGAGGCTCAATGGTCAAGTTCACGGCGCGGACCTTGGAACT 794
 QY 721 TATCTGCTGACGCTGTGCGCAACGGCGCGGACTCTACCGCATCCAGCATCTCTCAAC 780
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 795 TATCTGCTGACGCTGTGCGCAACGGCGCGGACTCTACCGCATCCAGCATCTCTCAAC 854
 QY 781 CCATCAACATCTTGTGTCAGGCTGTCTTTAGAGTCTGACTCCGGGGCCAAAG 840
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 855 CCATCAACATCTTGTGTCAGGCTGTCTTTAGAGTCTGACTCCGGGGCCAAAG 914
 QY 841 GTCACCGCAATGCGGCTGAGCTGCGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 915 GTCACCGCAATGCGGCTGAGCTGCGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
 QY 901 AAAGTGAAGTGAACAGCACCCGAGTCTGGGACTGCCATCTCTTACCAGGAGGAC 960
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 975 AAAGTGAAGTGAACAGCACCCGAGTCTGGGACTGCCATCTCTTACCAGGAGGAC 1034
 QY 961 CTGTGTGGAGCCACACCTGTGACACCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1035 CTGTGTGGAGCCACACCTGTGACACCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1094
 QY 1021 CCCAAGAGAGTCTGCTGCTCAATGAGGAGGATGGCTTCCATCAGCTTCCACTGCC 1080
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1095 CCCAAGAGAGTCTGCTGCTCAATGAGGAGGATGGCTTCCATCAGCTTCCACTGCC 1154
 QY 1081 CACGAGCTGGGCGCAGTGTTCACATGCCCCCATGAAAGTGTGAAAGTGTGTGAGGAGTG 1140
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1155 CACGAGCTGGGCGCAGTGTTCACATGCCCCCATGAAAGTGTGAAAGTGTGTGAGGAGTG 1214
 QY 1141 TTTGGGAAGCTCCGAGCCCAACCATGATGTCCCGACCTTCATCCAGATCAGCGTGGC 1200
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1215 TTTGGGAAGCTCCGAGCCCAACCATGATGTCCCGACCTTCATCCAGATCAGCGTGGC 1274
 QY 1201 AACCCCTGCTGAGCTGCAGTGTGCCATCATCCGACTTCCCTGGAGAGGCGACGCT 1260
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1275 AACCCCTGCTGAGCTGCAGTGTGCCATCATCCGACTTCCCTGGAGAGGCGACGCT 1334
 QY 1261 GACTGCTCTCTGGAGCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTCCCGGGGCGC 1320
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1335 GACTGCTCTCTGGAGCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTCCCGGGGCGC 1394

QY 1321 AGCTACACCTGAGCAGCAGTCCGAGCTGGCTTTTGGGCTGGCTCCCAAGCCCTGTCT 1380
Db 1395 AGCTACACCTGAGCAGCAGTCCGAGCTGGCTTTTGGGCTGGCTCCCAAGCCCTGTCT 1454
QY 1381 TACATGCACTGACCAAGCTGTGGTCCACCGGGAAGGCAAGGACATGTTGTGC 1440
Db 1455 TACATGCACTGACCAAGCTGTGGTCCACCGGGAAGGCAAGGACATGTTGTGC 1514
QY 1441 CAGACCGGCACATTCCTCGGCGGATGAGACAGCTGTGGGAGAGGCAAGCTGTGCCTC 1500
Db 1515 CAGACCGGCACATTCCTCGGCGGATGAGACAGCTGTGGGAGAGGCAAGCTGTGCCTC 1574
QY 1501 AAGGGGCTGCTGGGAGACACACCTCAACAACACAGAGTGTGATGTTCTCGGGGC 1560
Db 1575 AAGGGGCTGCTGGGAGACACACCTCAACAACACAGAGTGTGATGTTCTCGGGGC 1634
QY 1561 AAATGGATCCCTATGCCCCCTCTCGCGCACATGTGGTGGGGCGTGCAGCTGGCCAGG 1620
Db 1635 AAATGGATCCCTATGCCCCCTCTCGCGCACATGTGGTGGGGCGTGCAGCTGGCCAGG 1694
QY 1621 AGGAGTGCACCAACCCACCTGCGAAGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
Db 1695 AGGAGTGCACCAACCCACCTGCGAAGGGGCAAGTACTGCGAGGAGTGAAGGTG 1754
QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGACATGCAAGCTCCGGAACAGCTTCCGG 1740
Db 1755 AAATACCGATCTGCAATCTGGAGCCCTGCCCGACATGCAAGCTCCGGAACAGCTTCCGG 1814
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGGTACACCAAGCAAGCAACCGGCTCACTTCGCG 1800
Db 1815 GAGGAGCAGTGTGAGGCTTTCAACGGGTACACCAAGCAAGCAACCGGCTCACTTCGCG 1874
QY 1801 GTGCACTGGTGGCCCAAGTACTCGGGTGTCTCCCGGCAAGTCAAGCTCACTTCG 1860
Db 1875 GTGCACTGGTGGCCCAAGTACTCGGGTGTCTCCCGGCAAGTCAAGCTCACTTCG 1934
QY 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTCTGTGCGACCCCAAGTGTGGTGGACGCA 1917
Db 1935 CGAGCCAAATGGCACTGGCTACTTCTATGTCTGTGCGACCCCAAGTGTGGTGGACGCA 1994
QY 1918 CTGTGCTCTGCTGACTCCACTCGGCTGTGTCTCAAGGCAAGTCAAGGCTGGCTGT 1977
Db 1995 CTGTGCTCTGCTGACTCCACTCGGCTGTGTCTCAAGGCAAGTCAAGGCTGGCTGT 2054
QY 1978 GATGGAACTGGGCTCCAAAGAGATTCGCAAGTGTGGGTGTGGGGGAGACAAT 2037
Db 2055 GATGGAACTGGGCTCCAAAGAGATTCGCAAGTGTGGGTGTGGGGGAGACAAT 2114
QY 2038 AAGAGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGATGCTACAATTCGTG 2097
Db 2115 AAGAGCTGCAAGAAGGTGACTGGACTTTCACCAAGCCCATGATGCTACAATTCGTG 2174
QY 2098 GTGGCCATCCCGCAGCGGCTCAAGCATCGACATCCGCTAGCGCGTTACAAGGGCTG 2157
Db 2175 GTGGCCATCCCGCAGCGGCTCAAGCATCGACATCCGCGCGGTTACAAGGGCTG 2234
QY 2158 ATCGGGATGACAATCACTACCTGGCTCTGAAGAAGCAAGCAAGTCACTGCTCAAGGG 2217
Db 2235 ATCGGGATGACAATCACTACCTGGCTCTGAAGAAGCAAGCAAGTCACTGCTCAAGGG 2294
QY 2218 CATTTGCTGTGTGCGGGTGGAGCGGACCTGGTGTGAAGGCAAGTCTGCTGCGGTAC 2277
Db 2295 CATTTGCTGTGTGCGGGTGGAGCGGACCTGGTGTGAAGGCAAGTCTGCTGCGGTAC 2354
QY 2278 AGCGCACGGGCACAGCGGTGGAGAGCTTCAGAGCTTCGCGGCCATCTTGGAGCGCTG 2337
Db 2355 AGCGCACGGGCACAGCGGTGGAGAGCTTCAGAGCTTCGCGGCCATCTTGGAGCGCTG 2414
QY 2338 ACCGTGAGGTCTCTCGTGGGGAAGTACACACCGCGGGTCCGCTACTCTCTCTAT 2397
Db 2415 ACCGTGAGGTCTCTCGTGGGGAAGTACACACCGCGGGTCCGCTACTCTCTCTAT 2474

QY 2398 CTGCCAAAGAGCCTCGGAGGACAAAGTCTCTATCCCAAGGACCCCGG---GACCC 2454
Db 2475 CTGCCAAAGAGCCTCGGAGGACAAAGTCTCTATCCCGCCGACCCCGGGAGGACCC 2534
QY 2455 TCTGTCTTGCACAACAGCGTCTCAGCTCTCCAAACAGGTGGAGCAGCGGAGCAGG 2514
Db 2535 TCTGTCTTGCACAACAGCGTCTCAGCTCTCCAAACAGGTGGAGCAGCGGAGCAGG 2594
QY 2515 CCCCCTGCACCTGSGTGGCTGGCAGCTGGGGCGCTGCTCCGCGAGCTGCGGCAGTGGC 2574
Db 2595 CCCCCTGCACCTGSGTGGCTGGCAGCTGGGGCGCTGCTCCGCGAGCTGCGGCAGTGGC 2654
QY 2575 CTGAGAAAGCGGGCGGTGAGCTGTGGGGGTCCCGCGGAGCAGCGTCCCTGCCI3T 2634
Db 2655 CTGAGAAAGCGGGCGGTGAGCTGTGGGGGTCCCGCGGAGCAGCGTCCCTGCCI3T 2714
QY 2635 GATCAGCCCATCGGCCGCTGGAGACACAGCCTGCGGGGAGCCTGCCACCTGCG3AG 2694
Db 2715 GATCAGCCCATCGGCCGCTGGAGACACAGCCTGCGGGGAGCCTGCCACCTGCG3AG 2774
QY 2695 CTCAGCGCTGTGCTCACCCTGCTCCAAAGAGCTGCGGCGGGGATTTCAGAGCGGCTCACIC 2754
Db 2775 CTCAGCGCTGTGCTCACCCTGCTCCAAAGAGCTGCGGCGGGGATTTCAGAGCGGCTCACIC 2834
QY 2755 AAGTGTGGGCGCAGGAGCGCGCTGCTGCCCGGAGCAGTCAACTTGCACCGCAAG 2814
Db 2835 AAGTGTGGGCGCAGGAGCGCGCTGCTGCCCGGAGCAGTCAACTTGCACCGCAAG 2894
QY 2815 CCCAGGAGCTGGACTTCTGCTGCTGAGGCGGTGC 2850
Db 2895 CCCAGGAGCTGGACTTCTGCTGCTGAGGCGGTGC 2930
RESULT 5
AAS97182
ID AAS97182 standard; cdNA; 2937 BP.
XX AAS97182;
AC AAS97182;
DT 26-FEB-2002 (first entry)
XX Human metalloprotease partial DNA sequence #11.
DE Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;
KW vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;
KW hypertensive; neuroleptic; neuroprotective; anabolic;
KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;
KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
KW immune-related disease; cardiovascular disease; neuronal disease;
KW migraine; sexual dysfunction; mood disorder; attention disorder;
KW cognition disorder; hypotension; hypertension; psychotic disorder;
KW dyskinesia; metabolic disorder; inflammatory disorder; ss.
XX Homo sapiens.
XX WO200183782-A2.
PN 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US14431.
PF 04-MAY-2000; 2000US-201879P.
PR (SUG-) SUGEN INC.
XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI Payne V;
XX WPI; 2002-041502/05.
DR P-PSDB; AAU72899.
XX Novel protease polypeptide useful for screening for substances that may

be used to treat, e.g., cancers, immune-related diseases,
cardiovascular disease, migraine, pain, psychotic and inflammatory
disorders -

Claim 30; Figure 1P; 232pp; English.

The invention relates to an isolated, enriched, or purified protease polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to screen for substances (S) that may modulate its activity. Administering S (which modulates protease activity in vitro) may be used to treat a disease or disorder selected from cancers (e.g., of tissues, of blood or haematopoietic origin, of the breast, colon, lung, prostate, cervical, brain, ovarian, bladder or kidney), immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases (e.g., central or peripheral nervous system diseases, migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders and dyskinesias), metabolic disorders and inflammatory disorders. (I) may also be useful as a diagnostic tool for a disease or disorder such as those above. AAS97159-AAS97195 represent human protease coding sequences and primers of the invention.

Sequence 2937 BP; 531 A; 958 C; 936 G; 512 T; 0 other;

Query Match 98.8%; Score 2817.6; DB 24; Length 2937;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2849; Conservative 0; Mismatches 4; Indels 9; Gaps 2;

PT	1	ATGCTTCTGCTGGGATCTTAACCTCTGGCTTTCGCGGGGAGAACCGTGGAGGCTCTGAG	60
PT			
PS	76	ATGCTTCTGCTGGGATCTTAACCTCTGGCTTTCGCGGGGAGAACCGTGGAGGCTCTGAG	135
PS			
XX	61	CCAGAGGGGAGGAGTGGTTCCTCCATCCGACTGGACCCCGGACATTAAACGGCCGCGCTAC	120
XX			
CC	136	CCAGAGGGGAGGAGTGGTTCCTCCATCCGACTGGACCCCGGACATTAAACGGCCGCGCTAC	195
CC			
CC	121	TACTGGGGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	180
CC			
CC	196	TACTGGGGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	255
CC			
QY	181	CAGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG	240
QY			
DB	256	CAGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG	315
DB			
QY	241	ACTGAGCATGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG	300
QY			
DB	316	ACTGAGCATGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG	375
DB			
QY	301	TGCTTCTATTCTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	360
QY			
DB	376	TGCTTCTATTCTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	435
DB			
QY	361	GGGGGGTCCGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAG	420
QY			
DB	436	GGGGGGTCCGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAG	495
DB			
QY	421	AATGCTAGCGCGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	480
QY			
DB	496	AATGCTAGCGCGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	555
DB			
QY	481	GGTGTTCGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG	540
QY			
DB	556	GGTGTTCGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGG	615
DB			
QY	541	AACCCGCCATCTACGGGCGCTGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	600
QY			
DB	616	AACCCGCCATCTACGGGCGCTGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	675
DB			
QY	601	AGTGTAGCGCGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	660
QY			
DB	676	AGTGTAGCGCGCGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	735
DB			
QY	661	GAGAGCGTGGTGGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGT	720
QY			

Db	736	GAGAGCGTGGTGGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG	795
QY			
QY	721	TATCTGCTGACGCTGGTGGCAACGGCGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	780
QY			
Db	796	TATCTGCTGACGCTGGTGGCAACGGCGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	855
Db			
QY	781	CCCATCAACATCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	840
QY			
Db	856	CCCATCAACATCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	915
Db			
QY	841	GTACACGGCAATCGCGCCCTGGAGCGCTGGAGCGCTGGAGCGCTGGAGCGCTGGAGCGCTGGAG	900
QY			
Db	916	GTACACGGCAATCGCGCCCTGGAGCGCTGGAGCGCTGGAGCGCTGGAGCGCTGGAGCGCTGGAG	975
Db			
QY	901	AAAGTGAAGTGAACAGCAACCCCGAGTACTGGGACACTGGCCTCTCTTCCACAGGAGGAGGAG	960
QY			
Db	976	AAAGTGAAGTGAACAGCAACCCCGAGTACTGGGACACTGGCCTCTCTTCCACAGGAGGAGGAG	1035
Db			
QY	961	CTGTGTGGAGCCACCCTGTGACACCCCTGGGAGCGCTGGGAGCGCTGGGAGCGCTGGGAGCG	1020
QY			
Db	1036	CTGTGTGGAGCCACCCTGTGACACCCCTGGGAGCGCTGGGAGCGCTGGGAGCGCTGGGAGCG	1095
Db			
QY	1021	CCCAAGAGAGTGGTCTCTGCTGATTGAGGAGGAGTGGGCTTCCATCAGCTTCCACTTCCACTGCT	1080
QY			
Db	1096	CCCAAGAGAGTGGTCTCTGCTGATTGAGGAGGAGTGGGCTTCCATCAGCTTCCACTTCCACTGCT	1155
Db			
QY	1081	CAGGAGCTGGGCGACGCTGTTCAACATGCCCCATGACAAATGTGAAAGTCTGTGAGGAGGAG	1140
QY			
Db	1156	CAGGAGCTGGGCGACGCTGTTCAACATGCCCCATGACAAATGTGAAAGTCTGTGAGGAGGAG	1215
Db			
QY	1141	TTTGGGAAGCTCCGAGCCCAACCAACATGATGTCGCCAGCCCTCATCCAGATCACCAGTGGC	1200
QY			
Db	1216	TTTGGGAAGCTCCGAGCCCAACCAACATGATGTCGCCAGCCCTCATCCAGATCACCAGTGGC	1275
Db			
QY	1201	AACCCCTGTGACGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1260
QY			
Db	1276	AACCCCTGTGACGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1335
Db			
QY	1261	GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCGAGGAGTGGTGGGCGGCGC	1320
QY			
Db	1336	GACTGCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCGAGGAGTGGTGGGCGGCGC	1395
Db			
QY	1321	AGCTACACCTGAGCGAGCAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	1380
QY			
Db	1396	AGCTACACCTGAGCGAGCAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	1455
Db			
QY	1381	TACATGAGTACTGACCAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1440
QY			
Db	1456	TACATGAGTACTGACCAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1515
Db			
QY	1441	CAGACCGGCACTTCCCTGGGCGGATGGCACCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	1500
QY			
Db	1516	CAGACCGGCACTTCCCTGGGCGGATGGCACCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	1575
Db			
QY	1501	AAAGGGGCTGGTGGAGAGACAAACCTCAACAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1560
QY			
Db	1576	AAAGGGGCTGGTGGAGAGACAAACCTCAACAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1635
Db			
QY	1561	AAATGGGATCCCTATGGGCGGCTGGTGGGCGACATGGTGGGCGGCTGGGCGGCTGGGCGGCTGGG	1620
QY			
Db	1636	AAATGGGATCCCTATGGGCGGCTGGTGGGCGACATGGTGGGCGGCTGGGCGGCTGGGCGGCTGGG	1695
Db			
QY	1621	AGGAGTGGCACCACCCACCCCTGCCAACGGGGGCAAGTACTGCGAGGAGGAGTGGGAGGAGGAGG	1680
QY			
Db	1696	AGGAGTGGCACCACCCACCCCTGCCAACGGGGGCAAGTACTGCGAGGAGGAGTGGGAGGAGGAGG	1755
Db			
QY	1681	AAATACCAGTCTGCAATCTGGAGCGCTGCCCGAGCTCAGGCTCCGGAAGAGGCTTCCGG	1740
QY			
Db	1756	AAATACCAGTCTGCAATCTGGAGCGCTGCCCGAGCTCAGGCTCCGGAAGAGGCTTCCGG	1815
Db			
QY	1741	GAGGAGCAGTGTGAGGCTTTCACGGCTTAACACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1800
QY			

QY 325 GCCGAGCCGGACTCGCTTCGCTGCTGTGAGCCTGTGCGGGGGGCTCCGCGGAGCTTTGGC 384

Db 97 GCGAGCCAGACTCCTTTGCTGCTAAGCCTATGCGGGGTCTCCGCGAGCCTTTGGC 156
QY 385 TACCAGAGCCGCGAGTATGCTATAGCCCGCTGCCCAATGCTAGCGCGCCGGCGGCAG 444
Db 157 TACCAAGGTGGGAGTATGCTATTAGCCCTCTGCCCAACACAGCGCGCCTGAGGGCAG 216
QY 445 CGCAACACGAGGGCGCACACTTCTCCAGCGCCGGGTGTTCGGGCGGCGCTTCCGGA 504
Db 217 CGTATAGCCAGGGCGCACACTTCTCCAGCGCCGGGTGCTCCGCTAGGCGCTTCCGGA 276
QY 505 GACCCCACTCTCCGTGCGGGGTGGCTTCGGGTGGAAACCCCGCATCTTACGGGGCCCTG 564
Db 277 GACCCCTACCTCTCCGTGCGGGGTGGCTTCGGGTGGAAACCCCGCATCTTACGGGGCCCTG 336
QY 565 GACCCCTTACAAGCGCGGGCGGGCTTCGGGGAGAGTCTAGCCGCGCAGCTCTGGG 624
Db 337 GACCCCTTAAACACAGCGCGGAGCGGCTGGCGGAAGCCCAACCCGCGCAGGTCTGGG 396
QY 625 CGCGCCAGCGTTCGTCTATCCCGCGGTACGTGGAGCGCTGGTGGTGGGGAGAG 684
Db 397 CGCGCCAGCGTTCGTCTATACCAGGTACGTGGAGACACTGGTGGTGGCGGAGAG 456
QY 685 TCAATGTCAAGTTCACGGCGCGGACCTCGAACATTTATCTGCTGACGCTCTGGCAACG 744
Db 457 TCAATGTCAAGTTCACGGCGCGGATTGGAACTATCTGCTGACGCTCTGGCCACG 516
QY 745 GCGGCGGACTCTACCGGCATCCAGCATCTCAACCCCATCAACATCTGTTGTTCAAG 804
Db 517 GCGGCGGACTCTACCGGCACCCAGCATCTCAACCTATCAACATCTGTTGTTCAAG 576
QY 805 GTGCTGCTTTAGAGATCGTACTCCGGGCGCCAAAGTTCACGGCAATGGGCGCTGACG 864
Db 577 GTGTTACTTTAGGAGATCGTACACTGGGCGCAAGGTTCACAGGCAACGGCGCCTGACT 636
QY 865 CTGCGCACTCTCTGCTCGGCGAGAGAGCTGAACAAAGTGAAGCAACGCCCGAG 924
Db 637 CTGCGCACTCTCTGCTCGGCGAGAAAGTTGAACAAAGTGAAGCAACGCCCGAG 696
QY 925 TACTGGGACATGCCATCCTTCCAGGAGGACCTGTGTGGAGCCACACCTGTGAC 984
Db 697 TACTGGGACACAGCATCCTCTTCCAGGAGGACCTATGGGGGTACCACTGTGAC 756
QY 985 ACCCTGGGATGCGTATGCTGGTACCATCTGTGACCCCAAGACAGCTGCTCTCAT 1044
Db 757 ACCTGGGATGCTATGCTGGGACCATGTGTGATCCCAAGAGAGCTGCTGTGATC 816
QY 1045 GAGGAGATGGGCTTCATCAGCCTTCACTGCCCCAGAGCTGGGCGCATGTTCAAC 1104
Db 817 GAGGAGATGGGCTTCGCTCGGCTTCACTGCCCCAGAGCTGGGCGCATGTTCAAC 876
QY 1105 ATGCCCCATGACATGTGAAGTCTGTGAGGAGGTGTTGGAGACTCCGAGCCACAC 1164
Db 877 ATGCCCCATGACATGTGAAGTGTGTGAGGAGGTGTTGGAGACTCCGAGCCACAC 936
QY 1165 ATGATGTCCCGACCTCATCAGATCGACCTGCAACCCCTGGTCAAGCTGCAAGTGT 1224
Db 937 ATGATGTCCGACACTCATCAGATCGACCTGCAACCCCTGGTCAAGCTGCAAGTGT 996
QY 1225 GCATCATCAGGACTTCCTGGAGCGGGACGGTGAAGTCTGCTCTGGAGCAACCCAGC 1284
Db 997 GCATCATCAGGACTTCCTGGAGCGGGACGGTGAAGTGTGAGGAGGTGTTGGAGACTCCGAGCCACAC 1056
QY 1285 AAGCCATCTCCCTGCGGAGATCTGCGGGCGGACCTACACCTGAGCCAGCAGTGC 1344
Db 1057 AAGCCATCACCCTGCTGAGGACTTGCAGGCAAGCTAGAGTTGAGGCCACAGTGC 1116
QY 1345 GAGCTGGCTTTGGCGTGGCTTCAAGCCCTGTCTTACATGACTGACTGACCAAGCTG 1404
Db 1117 GAGCTGGCTTTGGCGTGGCTTCAAGCCCTGAGCCCTGCCCCATATGCACTACTGTACAAAGCTG 1176
QY 1405 TGTGTACCGGGAGCCCAAGGAGAGATGTTGTGCGACAGCCCGCACTTCCCTGGGCC 1464
Db 1177 TGTGTACCGGGAGCCCAAGGAGAGATGTTGTGCGACAGCTCGGCCACTTCCCTGGGCA 1236

QY 1465 GATGCCACACAGCTGTGGGAGGCAAGCTCTGCTCAAGGGGCTCGCTGGAGAGACAC 1524
Db 1237 GATGCCACACAGCTGTGGGAGGCAAGTCTGCTCAAGGGGCTCGCTGGAGAGACAC 1296
QY 1525 AACCTCAACAGCAGAGGTGGATGGTTCTTGGGCCAAATGGGATCCCTATGGCCCTGC 1584
Db 1297 AACCCAAACAGTACCGGGTGGACGGCCTTGGGCCAAGTGGGAGCCTTATGGTCCCTGC 1356
QY 1585 TCGCCACATGTGTGGGGGCTGCAGCTGGCAGGAGGCA--GTGCACCAACCCACCC 1642
Db 1357 TCGCCACCTGCTGGTGGGGGCGCAGCTGGCCGAGGCAAGTGCAGACACCCACCC 1416
QY 1643 CTGCCAACGGG--GGCAAGTACTGCGAGGAGTGAAGGTGAATACCGATCTCGCAATCTG 1701
Db 1417 CTGCCAACGGGCGGAGTACTGCGAGGAGTGAAGTGAATACCGATCTCGCAATCTG 1476
QY 1702 GAGCCCTGCCCGAGCTCAGCTCCGGAAGAGCTTCGCGGA 1742
Db 1477 GAGCCCTGCCCGAGCTCAGCTTCGCAAGAGCTTCGCGGA 1517

RESULT 7

AAD35570

ID AAD35570 standard; cDNA; 1104 BP.

XX AAD35570;

XX AC 26-JUL-2002 (first entry)

XX Human protease cDNA #3.

XX Human; novel human protein; NHP; protease; biological disorder; obesity;
KW high blood pressure; arthritis; connective tissue disorder; infertility;
KW gene therapy; enzyme; gene; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..1104

FT /tag= a

FT /product= "Human protease #3"

XX WO200226949-A2.

XX PD 04-APR-2002.

XX PF 27-SEP-2001; 2001WO-US30350.

XX PR 29-SEP-2000; 2000US-236689P.

XX (LEXI-) LEXICON GENETICS INC.

XX Friddle CJ, Hilbun E;

XX WPI; 2002-372123/40.

XX P-PSDB; AAE22542.

XX Novel nucleic acid encoding a human protease, useful as a hybridization
PT probe for screening libraries and assessing gene expression patterns -
PS Disclosure; Page 38-39; 41pp; English.

XX The present sequence is a cDNA encoding novel human protein (NHP).
CC human protease. NHPs share structural similarity with animal proteases
CC particularly zinc metalloproteases. Sequences of the invention are
CC useful in therapeutic, diagnostic and pharmacogenomic applications.
CC NHP polynucleotides are used as hybridisation probes for screening
CC libraries and assessing gene expression patterns. They can also be
CC used for treating related biological disorders such as obesity, high
CC blood pressure, arthritis, connective tissue disorders and infertility.
CC They are also used in gene therapy.

XX

SQ		Sequence 1104 BP; 189 A; 357 C; 345 G; 213 T; 0 other;			
Query Match		38.2%; Score 1091.2; DB 24; Length 1104;			
Best Local Similarity		99.7%; Pred. No. 2.3e-218;			
Matches 1093; Conservative		0; Mismatches 3; Indels 0; Gaps 0;			
Qy	1	ATGCTTCTGCTGGCATCTTAACCTTGGCTTTCGCGGCGCAACGCTGGAGCTCTGAG	60		
Db	1	ATGCTTCTGCTGGCATCTTAACCTTGGCTTTCGCGGCGCAACGCTGGAGCTCTGAG	60		
Qy	61	CCAGAGCGGAGTAGTCTTCCATCCGACTGGAGCCCGACATTAACGCGCGCGCTAC	120		
Db	61	CCAGAGCGGAGTAGTCTTCCATCCGACTGGAGCCCGACATTAACGCGCGCGCTAC	120		
Qy	121	TACTGGCGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCATTT	180		
Db	121	TACTGGCGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCATTT	180		
Qy	181	CAGGAGACTTTTACCTACACCTGACCGCGGATGCTTCTTGGCTCCGCGCTTCTCC	240		
Db	181	CAGGAGACTTTTACCTACACCTGACCGCGGATGCTTCTTGGCTCCGCGCTTCTCC	240		
Qy	241	ACTGAGCATCTGGCGCTTCCCTCCAGGGGCTCACCGGGGGCTTCACACCTGCGACGC	300		
Db	241	ACTGAGCATCTGGCGCTTCCCTCCAGGGGCTCACCGGGGGCTTCACACCTGCGACGC	300		
Qy	301	TGCTTCTATTCTGGGAGCTGACCGCGGAGCTCGTTCGCTGCTGTGAGCCTGTGC	360		
Db	301	TGCTTCTATTCTGGGAGCTGACCGCGGAGCTCGTTCGCTGCTGTGAGCCTGTGC	360		
Qy	361	GGGGGCTCCGCGAGCTTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC	420		
Db	361	GGGGGCTCCGCGAGCTTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC	420		
Qy	421	AATGCTAGCGCGCGGCGGCGAGCGACAGCAGCGGGCGCACACCTTCTCCAGCGCGG	480		
Db	421	AATGCTAGCGCGCGGCGGCGAGCGACAGCAGCGGGCGCACACCTTCTCCAGCGCGG	480		
Qy	481	GGTGTTCGGGCGGCGCTTCCGAGAGCCACCTCTCGCTGCGGGGTGGGCTCGGCTGG	540		
Db	481	GGTGTTCGGGCGGCGCTTCCGAGAGCCACCTCTCGCTGCGGGGTGGGCTCGGCTGG	540		
Qy	541	AACCCCGCATCTTACGGGCGCTGGACCTTACAAGCGCGGGCGGGCTTCGGGGAG	600		
Db	541	AACCCCGCATCTTACGGGCGCTGGACCTTACAAGCGCGGGCGGGCTTCGGGGAG	600		
Qy	601	AGTGTAGCGGCGAGCTTGGGCGCGCAAGCTTTCTGCTATCCCGCGGTACGTG	660		
Db	601	AGTGTAGCGGCGAGCTTGGGCGCGCAAGCTTTCTGCTATCCCGCGGTACGTG	660		
Qy	661	GAGAGCTGTGTGCGGCGAGTCAATGTTCAAGTTCACGCGCGGACCTTGGAAACAT	720		
Db	661	GAGAGCTGTGTGCGGCGAGTCAATGTTCAAGTTCACGCGCGGACCTTGGAAACAT	720		
Qy	721	TATCTGTAGCTGTGTGGCAACGCGCGGCGGACTTACCGCATCCACATCTCTCAAC	780		
Db	721	TATCTGTAGCTGTGTGGCAACGCGCGGCGGACTTACCGCATCCACATCTCTCAAC	780		
Qy	781	CCCATCAACATCTGTGTGCTCAAGTCTGCTTCTTACAGATGCTGACTCCGGGCGCAAG	840		
Db	781	CCCATCAACATCTGTGTGCTCAAGTCTGCTTCTTACAGATGCTGACTCCGGGCGCAAG	840		
Qy	841	GTCACCGCAATGGGCGCTGACCTGCGCAACTTCTGTGCTGGCAGAGAAAGCTGAAC	900		
Db	841	GTCACCGCAATGGGCGCTGACCTGCGCAACTTCTGTGCTGGCAGAGAAAGCTGAAC	900		
Qy	901	AAAGTGTAGCAGACCGCGAGTACTGGGACACTGCCATCTCTTCACGAGCAGGAC	960		
Db	901	AAAGTGTAGCAGACCGCGAGTACTGGGACACTGCCATCTCTTCACGAGCAGGAC	960		
Qy	961	CTGTGTGGAGCCACCTGTGACACCTTGGGCTGCTGATGTGGGTACCATGTGTGAC	1020		
Db	961	CTGTGTGGAGCCACCTGTGACACCTTGGGCTGCTGATGTGGGTACCATGTGTGAC	1020		

Qy	1021	CCCAAGAGAACTGCTCTGCTTCAATTGAGGACGATGGGCTTCATCAGCCTTCACCACTGCC	1080
Db	1021	CCCAAGAGAACTGCTCTGCTTCAATTGAGGACGATGGGCTTCATCAGCCTTCACCACTGCC	1080
Qy	1081	CACGAGCTGGGCCACG	1096
Db	1081	CACGAGCTGGGTAAGG	1096
RESULT 8			
AAA95831			
ID	AAA95831 standard; cDNA; 1143 BP.		
XX	AC	AAA95831;	
XX	AC	AAA95831;	
DT	23-FEB-2001 (first entry)		
XX	Human metalloproteinase ADAMTS-5 cDNA.		
DE	Human; ADAMTS-5; metalloproteinase; ADAM;		
KW	a disintegrin and metalloproteinase domain; thrombospondin domain;		
KW	vacuine; nootropic; neuroprotective; antiparkinsonian;		
KW	cerebroprotective; cytosolic; antiarthritic; immunosuppressive;		
KW	Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;		
KW	autoimmune disease; brain tumour; brain injury; ss.		
XX			
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FH	CDS	1..1143	
FT		/*tag= a	
FT		/partial	
FT		/product= "ADAMTS-5"	
XX		WO200053774-A2.	
PN		14-SEP-2000.	
XX		08-MAR-2000; 2000WO-US06237.	
PF		08-MAR-1999; 99US-0264585.	
XX		(NEUR-) NEUROCRINE BIOSCIENCES INC.	
PA	Kelner GS, Clark M, Maki RA;		
XX		WPI; 2000-594326/56.	
XX		P-PSDB; AAB21261.	
DR		Polynucleotide encoding novel members of a disintegrin,	
DR		metalloproteinase and thrombospondin domain protein family used to	
XX		prevent and treat Alzheimer's disease, cancer and autoimmune diseases	
PT		Claim 2; Fig 22; 129pp; English.	
PS		The present sequence encodes human metalloproteinase ADAMTS-5. The	
CC		ADAMTS family of proteins is closely related to the ADAM (A Disintegrin	
CC		and Metalloproteinase domain) family. Members of the ADAMTS family	
CC		contain a thrombospondin domain in addition to the disintegrin and	
CC		metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are	
CC		useful for the manufacture of medicaments for treating conditions	
CC		associated with neuroinflammation and/or neurodegeneration, such as	
CC		Alzheimer's disease, Parkinson's disease and stroke. They are also	
CC		useful for treating conditions associated with cell proliferation, cell	
CC		migration, inflammation and/or angiogenesis, such as cancer, arthritis	
CC		and autoimmune diseases. They can be used to treat patients afflicted	
CC		with an invasive tumour, a brain tumour or brain injury.	
XX		Sequence 1143 BP; 240 A; 362 C; 326 G; 209 T; 6 other;	
SQ			
Query Match 36.6%; Score 1043.6; DB 21; Length 1143;			
Best Local Similarity 96.5%; Pred. No. 2e-208;			

Matches 1084; Conservative 2; Mismatches 34; Indels 3; Gaps 2;

```
QY 738 GCGACGGGGGGGCGGAGCTTACCGCATCCCGAGCATCTCAACCCCATCAACATGTTGT 797
Db 21 GCGCGCGGGGCGAGGTATCTTTAAGCATCCCGAGCATCTCAACCCCATCAACATGTTGT 80
QY 798 GGTCAAGTGTGCTTCTTAGAGTCGTGACTCCGGGCGGCAAGTCAACCGCAATGCGGC 857
Db 81 GGTCAAGTGTGCTTCTTAGAGTCGTGACTCCGGGCGGCAAGTCAACCGCAATGCGGC 140
QY 858 CTTGACGCTGCGCAACTTCTTGTGCTGCGAGAAGTGAACAAAGTGAAGTGAACAAGCA 917
Db 141 CTTGACGCTGCGCAACTTCTTGTGCTGCGAGAAGTGAACAAAGTGAAGTGAACAAGCA 200
QY 918 CCCCAGTACTGGGACACTGCCATCTCTTACACAGGAGGACCTGTGTGAGCCACAC 977
Db 201 CCCCAGTACTGGGACACTGCCATCTCTTACACAGGAGGACCTGTGTGAGCCACAC 260
QY 978 CTGTGACACCTTGGGCATGGTGTGATGGGTACCATGTGTACCCCAAGAGAACTGCTC 1037
Db 261 CTGTGACACCTTGGGCATGGTGTGATGGGTACCATGTGTACCCCAAGAGAACTGCTC 320
QY 1038 TGTCAATTGAGGACGATGGGCTTCCATCAGCCTTCAACACTGCCACGAGTGGGCCACGT 1097
Db 321 TGTCAATTGAGGACGATGGGCTTCCATCAGCCTTCAACACTGCCACGAGTGGGCCACGT 380
QY 1098 GTTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTGTGGGAAGCTCCGAGC 1157
Db 381 GTTCAACATGCCCATGACATGTGAAGTCTGTGAGGAGTGTGGGAAGCTCCGAGC 440
QY 1158 CAACACATGATGTCCCGACCTCATCCAGATCGACGTCGCAACCCCTGGTCAGCCTG 1217
Db 441 CAACACATGATGTCCCGACCTCATCCAGATCGACGTCGCAACCCCTGGTCAGCCTG 500
QY 1218 CAGTCTGCCATCATCACCGACTTCTTGACAGCGGCGACGCTGCTCTCTGGACCA 1277
Db 501 CAGTCTGCCATCATCACCGACTTCTTGACAGCGGCGACGCTGCTCTCTGGACCA 560
QY 1278 ACCGAGCAAGCCCATCTCCCTGCCGAGGATCTCCGGGCGGCGACGCTACACCTTGACCA 1337
Db 561 ACCGAGCAAGCCCATCTCTCCCTGCCGAGGATCTCCGGGCGGCGACGCTACACCTTGACCA 620
QY 1338 CGATGCGAGTGGCTTTTGGCGTGGCTCCAGCCCTGCTTACATGAGTACTGCAC 1397
Db 621 CGATGCGAGTGGCTTTTGGCGTGGCTTCAAGCCCTGCTTACATGAGTACTGCAC 680
QY 1398 CAAGTGTGTGTCACCGGGAAGGCAAGGACAGATGCTGTGCAGACGCGCCACTTCCC 1457
Db 681 CAAGTGTGTGTCACCGGGAAGGCAAGGACAGATGCTGTGCAGACGCGCCACTTCCC 740
QY 1458 CTGGCGGATGGCACCAAGCTGTGGCGAGGGAAGCTCTGCTCTCAAAGGGCCCTGCGTGA 1517
Db 741 CTGGCGGATGGCACCAAGCTGTGGCGAGGGAAGTCTGCTCTCAAAGGGCCCTGCGTGA 800
QY 1518 GAGACAACTCAACAGCACAGGCTGATGCTGCTGGCCAAATGGATCCCTATGG 1577
Db 801 AARACAACTCAACAGCACAGGCTGATGCTGCTGGCCAAATGGATCCCTATGG 860
QY 1578 CCCCTGCTCGCGCACATGTGTGGGGCGGTGCAGCTGGCCAGGAGGAG-TGCACCAAC 1636
Db 861 CCCCTGCTCGCGCACATGTGTGGGGCGGTGCAGTGGCCAGGAGGAGTGCACCAAC 920
QY 1637 CCACCCC--TGCCAACGGGGCAAGTACTCGAGGGAGTGAAGTAAATACCGATCCTG 1694
Db 921 CCANCCCTGCAACNCGGGGCAAGTACTCGAGGGAGTGAAGTAAATACCGATCCTG 980
QY 1695 CAATCTGGCCCTGCCCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGA 1754
Db 981 CAACCTGGAGCCCTGCCCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGA 1040
QY 1755 GGCTTTTCAACGGGTACAAACACAGCACCAACCGGCTCAGTCTCGCGTGGCATGGTGC 1814
Db 1041 GGCTTTTCAACGGGTACAAACACAGCACCAACCGGCTCAGTCTCGCGTGGCATGGTGC 1100
```

```
QY 1815 CAAGTACTCGGGGTGTCTCCCGGGACAAAGTGAAGTCAAGTCTATC 1857
Db 1101 CAAGTACTCGGGGTGTCTCCCGGTGACAAGTGAAGTCTATC 1143
```

RESULT 9

AAD35568
ID AAD35568 standard; cDNA; 966 BP.

XX AAD35568;

XX AC AAD35568;

XX 26-JUL-2002 (first entry)

XX Human protease cDNA #1.

XX Human; novel human protein; NHP; protease; biological disorder; obesity;
XX high blood pressure; arthritis; connective tissue disorder; infertility;
XX gene therapy; enzyme; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..966

XX FT /*tag- a

XX FT /product= "Human protease #1"

XX WO200226949-A2.

XX PD 04-APR-2002.

XX PF 27-SEP-2001; 2001WO-US30350.

XX PR 29-SEP-2000; 2000US-236689P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Friddle CJ, Hilbun E;

XX DR WPI; 2002-372123/40.

XX DR P-PSDB; AAE22540.

XX Novel nucleic acid encoding a human protease, useful as a hybridization
XX probe for screening libraries and assessing gene expression patterns -
XX Disclosure; Page 34; 41pp; English.

XX The present sequence is a cDNA encoding novel human protein (NHP),
XX human protease. NHPs share structural similarity with animal proteases
XX particularly zinc metalloproteases. Sequences of the invention are
XX useful in therapeutic, diagnostic and pharmacogenomic applications.
XX NHP polynucleotides are used as hybridisation probes for screening
XX libraries and assessing gene expression patterns. They can also be
XX used for treating related biological disorders such as obesity, high
XX blood pressure, arthritis, connective tissue disorders and infertility.
XX They are also used in gene therapy.

SQ Sequence 966 BP; 161 A; 316 C; 305 G; 184 T; 0 other;

Query Match 33.6%; Score 959; DB 24; Length 966;

Best Local Similarity 100.0%; Pred. No. 8.8e-191;

Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTCGGCGATCCTTAACCTCGGTTTCGGCGGGGGAACCGCTGGAGGCTTCAG 60

Db 1 ATGCTTCTGCTCGGCGATCCTTAACCTCGGTTTCGGCGGGGGAACCGCTGGAGGCTTCAG 60

QY 61 CCAGAGCGGGAGGTAGTTCCTCCCATCCGACTGGACCGGACATTAACGGCGCCCTAC 120

Db 61 CCAGAGCGGGAGGTAGTTCCTCCCATCCGACTGGACCGGACATTAACGGCGCCCTAC 120

QY 121 TACTGCGGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180

Db 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCAGACATTT 180
Qy 181 CAGGAGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGCGCTCCCGCTTCTCC 240
Db 181 CAGGAGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTGCGCTCCCGCTTCTCC 240
Qy 241 ACTGAGCATCTGGCGCTCCCGCTCCAGGGCTCACCGGGGCTTTCAGACCTGCGACGC 300
Db 241 ACTGAGCATCTGGCGCTCCCGCTCCAGGGCTCACCGGGGCTTTCAGACCTGCGACGC 300
Qy 301 TGCTTCTATTCTGGGACGTGAACCGCGGAGCTCGTTCGCTGCTGTGAGCCCTGTC 360
Db 301 TGCTTCTATTCTGGGACGTGAACCGCGGAGCTCGTTCGCTGCTGTGAGCCCTGTC 360
Qy 361 GGGGGGCTCCCGAGCTTTGGCTACCGAGCGCGGAGTATGCTATGATGACCGCGCTGCC 420
Db 361 GGGGGGCTCCCGAGCTTTGGCTACCGAGCGCGGAGTATGCTATGATGACCGCGCTGCC 420
Qy 421 AATGCTAGCGCGCGCGGCGAGCGCAACAGCAGCGCGCACACCTTCTCCAGCGCGG 480
Db 421 AATGCTAGCGCGCGCGGCGAGCGCAACAGCAGCGCGCACACCTTCTCCAGCGCGG 480
Qy 481 GGTGTTCGGCGCGCGCTTCGGGAGACCCACCTCTCGCTCGCGGGTGGCTCGGGCTG 540
Db 481 GGTGTTCGGCGCGCGCTTCGGGAGACCCACCTCTCGCTCGCGGGTGGCTCGGGCTG 540
Qy 541 AACCCCGCATCTACGGGCTTCGACCTTACAGCGCGCGGGGCGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTACGGGCTTCGACCTTACAGCGCGCGGGGCGGCTTCGGGGAG 600
Qy 601 AGTCGTAGCGCGCGAGGTTCGGCGCGCGCAAGCGTTTCTGTATATCTGCGGCTGAGTG 660
Db 601 AGTCGTAGCGCGCGAGGTTCGGCGCGCGCAAGCGTTTCTGTATATCTGCGGCTGAGTG 660
Qy 661 GAGAGCTGTGTTCGGGAGAGTCAATGTCAAGTTCCAGCGCGGAGCTTGGAAAT 720
Db 661 GAGAGCTGTGTTCGGGAGAGTCAATGTCAAGTTCCAGCGCGGAGCTTGGAAAT 720
Qy 721 TATCTGTGAGCTGTGGCAACGCGCGGAGTCTACGCGCATCCAGCATCTCTCAAC 780
Db 721 TATCTGTGAGCTGTGGCAACGCGCGGAGTCTACGCGCATCCAGCATCTCTCAAC 780
Qy 781 CCATCATCATCTGTGTGAGTCAAGTGTCTTCTTATAGATCTGACTCGGGGCCAAG 840
Db 781 CCATCATCATCTGTGTGAGTCAAGTGTCTTCTTATAGATCTGACTCGGGGCCAAG 840
Qy 841 GTCACCGCAATGGGCGCTGACGCTGCGCAACTTCTGTGCTGCGCAGAGAAAGCTGAAC 900
Db 841 GTCACCGCAATGGGCGCTGACGCTGCGCAACTTCTGTGCTGCGCAGAGAAAGCTGAAC 900
Qy 901 AAAGTGTGTGACAGCAACCGAGTACTGGGACACTGCGCATCTCTTTCACAGCAGGA 959
Db 901 AAAGTGTGTGACAGCAACCGAGTACTGGGACACTGCGCATCTCTTTCACAGCAGGA 959

RESULT 10

AAH20226

ID AAH20226 standard; DNA; 2670 BP.

XX

AC AAH20226;

XX

DT 07-AUG-2001 (first entry)

XX

DE Human ADAM-type metalloprotease-related probe, SEQ ID NO:3.

XX

KW Human; MDTs4; ADAM-type metalloprotease; drug screening;

XX

KW A Disintegrin And Metalloprotease; cancer; arthritis; probe; ss.

XX

OS Homo sapiens.

XX

PN JP2001017183-A.

XX

PD 23-JAN-2001.

XX 09-JUL-1999; 99JP-0196584.
XX 09-JUL-1999; 99JP-0196584.
PR (YAMA) YAMANOUCHI PHARM CO LTD.
XX WPI; 2001-275950/29.
XX A new metal protease and its preparation for use as an anti-cancer and
PT anti-arthritis therapeutic -
XX Example 1; Page 11-12; 22pp; Japanese.
XX The invention relates to the novel human ADAM (A Disintegrin And
CC Metalloprotease)-type metalloproteases MDTs4 (AAB73549) and MDTs5
CC (AAB73550). The metalloproteases can be used for the treatment of
CC cancers and arthritis. The invention also relates to the genes encoding
CC MDTs4 and MDTs5, vectors and host cells containing the MDTs4 or MDTs5
CC genes, the recombinant production of MDTs4 and MDTs5, and antibody
CC specific for MDTs4 or MDTs5, and methods of screening for compounds
CC which modulate the activity of MDTs4 and/or MDTs5. The present
CC sequence represents a human ADAM-type metalloprotease-related cDNA
CC sequence, used in an exemplification as a probe for screening human cDNA
CC libraries in order to isolate MDTs4 cDNA.
XX
SQ Sequence 2670 BP; 511 A; 853 C; 867 G; 439 T; 0 other;

Query Match 23.3%; Score 664.4; DB 22; Length 2670;
Best Local Similarity 59.2%; Pred. No. 3.3e-129;
Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;
Qy 157 CTCATTTTTCAGATCACAGCATTTTCAGGAGGACTTTTACCTACACTGACCGCGATCT 216
Db 157 CTCGCGCTCCACCTGTCGCCCTTCGCGCAAGGCTTCTGCTCGCGCTCGCGCCGAC 216
Qy 217 CAGTTCTTGGCTCCGCGCTTCTCCACTGAGCATCTGGCGTCCCGCTCCAGGGCTCACC 276
Db 217 AGCTTCTTGGCGCGGCTTCAAGATCGAGCGCTTCGGGGGCTTCGGCGGGCG---ACC 273
Qy 277 GGGGGGCTTTCAGACCTCGGACGCTGCTTCTATTCTGGGAGCTGAACCGCGAGCGGAC 336
Db 274 GGGGGGCGAGCGGGGCTCGCGGCTGCTTCTTCTCGGCACCGTGAATGGGAGCGCGAG 333
Qy 337 TCGTTCGCTGTGTGAGCTGTGCGGGGGGCTCCGCGGAGCTTTGGCTACCGAGCGGCC 396
Db 334 TCGTTCGCGCGGCTGAGCCTGTGCGCGGCTGAGCGGCTCCTTCTGCTGAGCGCGGAG 393
Qy 397 GAGTATGTCTATTAGCGCGCTGCGCAATGTAGCGCGCGCGCGCGCGAGCGCAACAGCG 456
Db 394 GAGTTCCACATCCAGCGCGAGCGCGGGGGCTCCCTGGCTCAGCGCGACCGCTGCAG 453
Qy 457 GCGGCACACCTTCTCCAGCGCGGGGTGTTCGCGGGGCGCTTCGCGGAGACCCACCTCT 516
Db 454 CGCTGGGGTCCGCGCGAGCGCGCGCTTCGCGCGAGGACCGCGAGTGGGAGTGGAGAG 513
Qy 517 CGCTCGGGGTG-GCCTCGGGCTGGNAACCGCGCATCTTACGGCGCTCGACCTTACAA 575
Db 514 GGAGAGGGTTCAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
Qy 576 GCGCGCGCGCGGGGCTTCGGGGAGAGTCTGAGCGCGCGCAGG- ----TCTGGCGCGCC 630
Db 574 GAAGAGAGGAG 633
Qy 631 AAGCGTTTCTGTCTATCCCGGGTACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 690
Db 634 AAGCGTTTCTGTCTATCCCGGGTTCGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAG 693
Qy 691 GTCAAGTTCACGCGCGGAGCTGGAACATTTCTCTGACGCTGCTGCTGCAACAGCGGGG 750
Db 694 GCTGCTTCTACGGGGCGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 753
Qy 751 CGACTCTACGGCATCCAGCATCTTCAACCCCATCAACATCGTTGTGTGTCAGAGTGTG 810

Db	754	CGAATCTACAAAGCACCAGCATCAAGAATTCCAATCAACCTGATGTTGGTGTAAAGTGCCTG	813
QY	811	CTTCTTAGAGATCGTGACTTCGGGGCCCAAGGTCAACGGCAATCGGCCCTTGACGCTGC	870
Db	814	ATCGTAGAAGATGAATAATGGGCCAGAGGTGCCACAAATGGGGGCTTACACTGCCT	873
QY	871	AACCTCTGTGCTGGCAGAGAGACTGAACAAGTGTAGTACAGACACCACCGAGTACTGG	930
Db	874	AACCTTGTCAACTGCGACGCGGCTTTCAACAGCGCCAGCAGCCGCCACAGAGCACTAC	933
QY	931	GACACTGCCATCCTCTTCCACAGCGAGGACCTGTGTGG---AGCCACCACCTGTGACACC	987
Db	934	GACACGGCATCTCTCTTACCAGACAGAACTCTGTGGCAGGAGGGGCTGTGTGACACC	993
QY	988	CTGGCATGGCTGATGTGGGTACCATTGTGACCCCAAGAGAAAGTCTCTGTGCAATTGAG	1047
Db	994	CTGGGTGTGGCAGACATCGGACCAATTTGTACCCCAACAAGAGCTCTCTCGTGTATCGAG	1053
QY	1048	GAGATGGGCTTCCATAGCCTTACCACCTTCCACAGAGCTGGGCCACGTTTCAACATG	1107
Db	1054	GATGAGGGGCTCCAGCGGCCCAACACCTTGGCCCATGAATTAGGGACAGTCTCTCAGCATG	1113
QY	1108	CCCATGCAATGTGAAGTCTGTGAGGAGTGTTTGGGAAGTCCGAGTCAACACATG	1167
Db	1114	CCCCAGCAGACTCCAGGCCCTGCACACGGCTCTTCTGGGCCCATGGCAAGCACAGTG	1173
QY	1168	ATGTCCTCCGACCCCTCATTCAGATTCAGCGTGCACACCTCTGTGTGAGCTTGCAGTGTGCC	1227
Db	1174	ATGGCACCGCTGTGTGCTCCACTGAACAGACGCTGCGCTTGGTCCCTCTCAGCGCCATG	1233
QY	1228	ATCATCACGCACTTCTGTGACAGCGGCAAGTGAAGTCTGCCCTCTGACCAAGCCAGCAAG	1287
Db	1234	TATCTCACAGAGCTTCTGGACGGCGGGCAGGAGACTGTCTCTTGATGCCCTCTGTGCG	1293
QY	1288	CCCATCTCCTGCCGAGGATTCGCGGGC-----GCCAGCTACACCTTGACCCAGCAG	1341
Db	1294	GCCTTGCCTTCCCAACAGCGCTCCCGGGCGCAGTGGCCCTGTACCAGCTGGACCAAG	1353
QY	1342	TGCGAGCTGGCTTTTGGCTGGGCTTCCAGCCGCTGTCTTACA-----TGCACTAC	1392
Db	1354	TGCAGCAGATCTTTGGGCGGATTTCCGCCACTTGCCTCCCAACACCTCTGTCTAGGAGCTC	1413
QY	1393	TGCACCAAGCTGTGTG---CACCGGGAAGCCAGGACAGATGGTGTGCCACACCCGC	1449
Db	1414	TGCGCCACGCTTTGTTGCCACACTGATGGGCTGAGCCCTGTGCCACAGAAGATGGC	1473
QY	1450	CACCTTCCCTGGCCGATGGCAGCAGCTGTGGCAGGCGCAGCTCTGCCCAAGGGGCC	1509
Db	1474	AGCCTGCCCTTGGCTTCACGGCAAGCCGTGCGGGCCCTGGGCACTTCTGTAGAGGAGC	1533
QY	1510	TG-----CGTGAGAGACACAACCTCAACAGACAGGTTGGATGTCTCTGGGCCAAA	1563
Db	1534	TGCTACCTGAGSAGGAAGTGGAGAGGCCCAAGCCGCTGTAGATGAGGCTGGGCCACCG	1593
QY	1564	TGGATTCCTATGCCCCCTGTCTGGGCACATGTGTGGGGCGGTGCAGCTTGGCCAGGAGG	1623
Db	1594	TGGGAGACCTTGGGGAAGATTTCTCGGACCTGTGGAGGAGGATACAGTTTTCACACCGT	1653
QY	1624	CAGTGCACCAACCCCAACCTTCCCAACGGGGCAAGTACTGCGAGGGAGTTCAGGGTGA	1683
Db	1654	GAGTGCAGGACCCGAGCCTCAGAAATGGAGAGATACTGCTTGGGTTCGAGAGGCCAAG	1713
QY	1684	TACCGATCTGCAATCTGAGGACCTTGCCTCCAGCTCAGGCTCCGGAAGAGCTTCCGGGAG	1743
Db	1714	TACAGTCTATGCCACACGAGGAATGCCCC-----CCTGACGGGAAAGATTCAGGGAG	1767
QY	1744	GAGCAGTGTGAGGCTTTCAACGCTTACACACACAGCACAACCGCTCACTCTCGCGCTG	1803
Db	1768	CAGCAGTGTGAGAAGTATTAATGCTTCAAAATTAATGATGAGCAGGGAATCT---CCTG	1824
QY	1804	GCATGGTGGCCCAAGTACTCGGCGCTGTCTCCCGGGGACAAGTCAAGCTCATCTGCCGA	1863

Db	1825	CAGTGGGTCCCAAGTATCTGGGTGTCTCCCGCGGACCCCTGCAGATTGTTCTCTGC	CGA	1888
Qy	1864	GCCAATTGGCACTGGCTACTTCTATGTGTGGGCACCCCAAGGTGGTGGAGCGCACGCTGTGC	1923	
Db	1885	GCCGGGGAGGAGCGAGTTCAAAGTGTTCGAGGCCAAAGGTGATGTATGGCACCCCTTG	1944	
Qy	1924	TCCTCTGACTCCACCTCCGCTCTGTGTCCAAAGGCAAGTGCATCAGGCTGGCTGTGATGGG	1983	
Db	1945	GGGCGGAGAAACACTGGCCATCTGTGTCCGTGGCCAGTGTGTCAAGCCGGCTGTGACCAT	2004	
Qy	1984	AACCTGGGCTCCCAAGAGAAGATTCGACAAAGTGTGGGGTGTGTGGGGGAGACAATAAGAGC	2043	
Db	2005	GTGGTGGACTCGCTCGGAAGCTGGACAAAATCGGGGTGTGTGGGGGCAAAAGGCAACTCC	2064	
Qy	2044	TGCAGAGAGTGACTGGACTCTTCCACCAAGCCCCATGCGATGGCTACAAATTTCTGTGGCC	2103	
Db	2065	TGCAGGAAGTCTCCGGGTCCCTTCACCCGCCCAATTTGGCTACATGACATTTGTCACC	2124	
Qy	2104	ATCCCCGAGCGGCTCAAGCATTCGACATCCGCCAGCGGGTTACAAAGGCTGATCGGG	2163	
Db	2125	ATCCCCAGCTGGTGCACATAATTGACGTGAAGCAGCGGAGCCACCGGGTGTGCAGAAC	2184	
Qy	2164	GATGACAACCTACTGGCTGTGAGAACACGCCAAGGCAAGTACCTGCTCAACGGGCAATTC	2223	
Db	2185	GATGGGAACCTACCTGGCGCTGAAGACGGCTGATGGGSCAGTACCTGCTCAACGGGCAACCTG	2244	
Qy	2224	GTGGTGTCCGGGTGGAGGGGACCTGGTGGTGAAGGGCAGTCTGCTCGGTTACAGCGGC	2283	
Db	2245	GCCATCTCTGCCATAGACAGGACATCTTGTGTGAAGGGGACCATCTCTGAATACAGCGGC	2304	
Qy	2284	ACGGGCACAGCGGTGGAGAGCCTGCAGGCTTCCCGGCCCATCTCTGGAGCCGCTGACCGTG	2343	
Db	2305	TCCATCGCCACCTTGGAGCGCTTCAGAGCTTCGGGCGCTTCCAGAGAGCTCTGACAGTG	2364	
Qy	2344	GAGTCTCTCTCCGT--GGGGAAGATGACACCGCCCCGGGTCCGTACTCTCTATCTG	2400	
Db	2365	CAGCTCTGACAGTCCCTGCGAGGCTTCTCCCGGCCCAAAAGTCAAAATACACCTTCTTTGTT	2424	
Qy	2401	CCCAAGA 2408		
Db	2425	CCTAATGA 2432		
RESULT 11				
AAF82166				
ID	AAF82166 standard; cdna; 2670 bp.			
XX	AAF82166;			
AC				
XX	02-JUL-2001 (first entry)			
DT	Human ADAM type metal. protease MDR53 encoding cdna SEQ ID NO:21.			
XX				
KW	Human; a disintegrin and metalloprotease type metal protease; MDR53;			
KW	MDR52; MDR53; ADAM type metal protease; cytosolic; antithrombotic;			
KW	cancer; arthritis; arthrosis deformans; ss.			
XX				
OS	Homo sapiens.			
XX				
PH	Key			
FT	CDS			
FT	Location/Qualifiers			
FT	1..2670			
FT	/*tag= a			
FT	/product= "MDR53"			
FT	/note= "a disintegrin and metalloprotease (ADAM)			
FT	type metal protease"			
XX				
PN	JP2001008687-A.			
XX				
PD	16-JAN-2001.			
XX				
PF	25-JUN-1999; 99JP-0180973.			
XX				
PR	25-JUN-1999; 99JP-0180973.			


```
xx (YAMA ) YAMANOUCHI PHARM CO LTD.
xx WPI: 2001-285362/30.
xx P-PSDB; AAB/4946.
xx New metal protease and metal protease gene, for use as a drug for
xx cancers, arthritis and arthrosis deformans
xx Example 3; Page 27-28; 31pp; Japanese.
xx The present sequence encodes a disintegrin and metalloprotease (ADAM)
xx type metal protease designated MDT83, isolated from human. MDT83 proteins
xx have cytostatic and antiarthritic activities. They can be used as a drug
xx for cancers, arthritis and arthrosis deformans.
xx
xx Sequence 2670 BP; 511 A; 853 C; 867 G; 439 T; 0 other;
xx
Query Match 23.3%; Score 664.4; DB 22; Length 2670;
Best Local Similarity 59.2%; Pred. No. 3.3e-129;
Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;
QY 157 CTCATTTTCAGATCAGACGATTCAGGAGGACTTTTACCTACACCTGACGCGGATGCT 216
DB 157 CTCGCGCTCCACCTGTCCGCTTCGGAAGGCTTCGTGCTGGCCCTGGGCGCCGACGAC 216
QY 217 CAGTTCTTGGCTCCGCGCTTCTCCACTGAGCATCTGGGCGTCCGCTCCAGGGGCTCAC 276
DB 217 AGCTTCCTGGCGCCGAGTTCAAGATCGAGCGCTCGGGGCTCCGCGCGGGCG---ACC 273
QY 277 GGGGGCTTCTCAGACCTGCGACGCTGCTTCTATTTCTGGGAGCTGACGCGGACGGAC 336
DB 274 GGGGGCGAGCGGGGCTGCGCGGCTGCTTCTTCTCCGGCACCTGTAATGGGGAGCCCGAG 333
QY 337 TCGTTCGCTGTGAGCTGTGGGGGGCTCCGCGAGCTTGGCTACCGAGGCGCC 396
DB 334 TCGCTGGCGGGCTGACCTGTGCGGGGCTGAGCGGCTTCTTCTGCTGGAGCGGAG 393
QY 397 GAGTATGTCATTTAGCCGCTGCCAATGCTAGCGCGCGCGGCGAGCGCAACAGCCAG 456
DB 394 GAGTTCACCATCCAGCGCGGCGGGGCTCCCTGCTCAGCGCGACCGCTGCGAG 453
QY 457 GCGGCACACCTTCTCAGCGCGGGGCTGTCGCGGGCGGGCTTCCGAGAGACCCACCTCT 516
DB 454 CGCTGGGCTCCCGCGGAGCGCGCCCTCCCGGAGGACCCGAGTGGGAGGTGGAGCG 513
QY 517 CGCTCGGGGTG-GCCTCGGCTGGAACCGCGCATCTCCTCGGSCCTGGACCTTACAA 575
DB 514 GGAGAGGTCAGAGGAGGAGAGGAGACACCGAGGAGGAGGAGAGGCCAA 573
QY 576 GCCGCGCGGGGCTTCGGGGAGAGTCGTAGCGCGCGCAGG-----TCTGGGCGCGCC 630
DB 574 GAAGAGGAGGAGGAGGCGCTAGCGAGCGCGCACCGCCCTGGGGCGCACGAGTAGGACC 633
QY 631 AAGCGTTTCGTCTATCCCGGCTAGTGGAGAGCTGTGTGTCGGGCGAGGAGTCAATG 690
DB 634 AAGCGTTTGTGTCTGAGCGCGCTGCTGGAAGACGCTGTGTGTGGCGGATGCTCCATG 693
QY 691 GTCAGTTCCAGCGCGGAGCTGGAACATTTATCTGTGAGCGCTGTGTGCAACGCGCGCG 750
DB 694 GCTGCTTCTACGGGCGAGCTCGAGAACCATCTCTGAGCTTAATGCTGTGGGAGCC 753
QY 751 CGACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCGTTGTGTGCAAGGTGCTG 810
DB 754 CGAATCTACAAGCACCCAGCATCAAGAAATTCATCAACCTGATGTGTGTTAAAGTCTG 813
QY 811 CTTCTTAGAGATGCTGACTCCGGGCCCAAGTCAACCGCATCGGCGCTGACGCTCGC 870
DB 814 ATCGTAGAAGATGAAAATGGGGCCAGAGGTGTCCGACATGGGGGGCTTACACTCGT 873
QY 871 AACTTCTGTGCTGCGCAGAAAGCTGAACAAAGTGAAGTGAACAGCACCGCCGCTACTGG 930
DB 874 AACTTCTGCACTGGCAGCGGCTTTCAACCGAGCCCGGAGCCGCCAGGAGCACTAC 933
```

```
QY 931 GACACTGCCATCCTCTTCCACGAGCAGGACCTGTGTG---AGCCACACCTGTGACACC 987
DB 934 GACAGCGCCATCCTGCTCACCAGACAGAACTTCTGTGGGACAGGGGCTGTGTGACACC 993
QY 988 CTGGCATGGCTGATGTGGGTACCATGTGTACCCCAAGAGAACTGCTCTGTCTCATTCAG 1047
DB 994 CTGGGTGTGGAGACATCGGGACCATTTTGTGACCCCAACAAAGCTGCTCCGCTGATCGAG 1053
QY 1048 GAGATGGGCTTCCATCAGCCTTCACCACTGCCACAGAGCTGGGCCACGCTTTCACCAATG 1107
DB 1054 GATGAGGGCTTCCAGGCGGCCACACACCTGSCCATGACTAGGCGACGCTCTCAGCATG 1113
QY 1108 CCCCATGCAATATGAAAGTGTGTGAGAGGTGTTTGGGAAGCTCCGAGCCAAACCAATG 1167
DB 1114 CCCCACGAGCACTCCAAAGCCCTGCACAGGCTCTTGGGGCCCATGGGCAAGCAGAGTG 1173
QY 1168 ATGTCCCGACCTTCATCCAGATCGACCGTGCACACCTGCTGTCAGCTGCAGTGTGCC 1227
DB 1174 ATGGCACCGCTTGTGCTCCACTGAACAGAGCTGCGCCCTGTGCTCCCTGAGCGCCATG 1233
QY 1228 ATCATCAGGCTTCTTGGACAGGGGCGACGCTGACTGCTCTCTGGACCAACCCAGCAAG 1287
DB 1234 TATCTCAGAGCTTCTTGACGGGGCGACGAGACTGTCTCTGGATGCTGCTGCTGCG 1293
QY 1288 CCAATCTCCTGCGCGAGGATCTGCCGGG-----GCCAGCTACACCTGAGCCAGCAG 1341
DB 1294 GCGCTGCGCCCTCCCGACAGGCTCCCGGGCGCATGGGCCCTGTACAGCTGGAGCAACAG 1353
QY 1342 TCGGAGCTGGCTTTTGGCTGCCAAGCCCTGCTCTTACA-----TGCAGTAC 1392
DB 1354 TGCAGCAGATCTTTGGCGCGGATTTCCGCCACTGCCCCAACACCTCTGCTCAGCAATC 1413
QY 1393 TGCACCAAGCTGTGTG---CACCGGGAAGCCAAAGGACAGATGGTGTGCCAGAG---GC 1449
DB 1414 TCGCGCCAGCTTTGTGTGCCACACTGATGGGCTGAGCGCCCTGTGCCACACGAAGAAFGC 1473
QY 1450 CACTTCCCTGCGCGATGGCAGCAGCTGTGCGAGGCAAGCTCTGCTCAAGAGGGGCC 1509
DB 1474 AGCCTGCGCTGGGCTGACGGCACCGCTGCGGGCTTGGGCGCTCTGCTCAGAGGAGC 1533
QY 1510 TG-----CGTGGAGAGACACAACTCAACAGCACAGGGTGGATGGTTCTTGGGCCAAA 1563
DB 1534 TGTCTACCTAGAGAGAACTGGAGAGGCCAAGCCGTTGATAGTGGAGCTGGSCACCG 1593
QY 1564 TGGGATTCCTATGGCCCTGCTCGCGCAATGTGTGGGGGCTGCGAGCTGGCCAGGAGG 1623
DB 1594 TGGGAGCCCTGGGAGAAATGTTCTCGGACCTGTGGAGGAGGAGTACAGTTTTCACACCGT 1653
QY 1624 CAGTGCACCAACCCACCCCTCCCAACGGGGCAAGTACTCGGAGGAGTGGAGGTGAAA 1683
DB 1654 GAGTGCAGAGAACCCGAGCGCTCAGAAATGGAGGAAGATACTGCTGGGTGGAGAGCCCAAG 1713
QY 1684 TACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCGCTCCGGAAGAGCTTCCGGGAG 1743
DB 1714 TACCACTATGCCACACAGGAAATGCCCG-----CCTGACGGGAAAGCTTTCAGGAG 1767
QY 1744 GAGCAGTGTGAGGCTTTCACCGGCTTACACACAGCAGACCAACCGGCTCACTCTCCCGTG 1803
DB 1768 CAGCAGTGTGAGAAGTATTAATGCTACAAATTAACCTGACATGAGCGGGAATCT---CCTG 1824
QY 1804 GCATGGTGGCCCAAGTACTCCCGGCTGTCTCCCGGACAAAGTCAAGCTCATCTGCGCA 1863
DB 1825 CAGTGGGTCCCAAGTATGCTGGGTGTCTCCCGGAGCCGCTGCAAGTGTGTCTGCGCA 1884
QY 1864 GCAATGTCACCTGGCTTACTTCTATGTGTGGCACCACCAAGTGTGGAGCGGACGCTGTGC 1923
DB 1885 GCCCGGGGAGGAGCGAGTTCAAAGTGTTCAGGCGCAAGTGATTGATGGCACCTGTGT 1944
QY 1924 TCTCCTGATCCACCTTCCCTGTGTGTCAGAGGCAAGTGCATCAAGCTGGCTGTGATGGG 1983
DB 1945 GGGCCAGAAACACTGGCCCATCTGTGTCCGTGGCCAGTGTGTCAAGGCGCGGCTGTGACCAT 2004
```


Db 817 ATCTAGAAAGTGAATAATGGGCCCCAGAGGTGTCGACAAATGGGGGCTTACACTGCGT 876
QY 871 AACTTCTGTCCTGGCAGAGAGAGCTGAACAAAGTGAAGTACACAAAGCACCAGTACTGG 930
Db 877 AACTTCTGCAACTGGCAGCGGCTTTCAACAGCCAGCGAGCGCCACCCAGAGCACTAC 936
QY 931 GACACTGCCATCTCTTACACAGGAGGACCTGTGTGG--AGCCACCACCTGTGACACC 987
Db 937 GACAGGCCATCTCTACACACAGAACTCTGTGGCAGGAGGGCTGTGTGACACC 996
QY 988 CTGGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGTCTGTGATTTAG 1047
Db 997 CTGGGTGTGCAGACATCGGAGCAATTTGTACCCCAACAAAAGCTGCTCGGTGATCGAG 1056
QY 1048 GACATGGGCTTCCATCAGCTTCCACACTGCCCAGAGCTGGCCAGCTGTGTTCACATG 1107
Db 1057 GATAGGGGCTCCAGGGGCCCCACCTGGCCCCATGAATAGGACACCTCTCTGATG 1116
QY 1108 CCCCATGACAAATGTGAAGTCTGTGAGAGGTGTGTGGGAAGCTCCGAGGCAACACCATG 1167
Db 1117 CCCACAGCACTCCAGACCTTGCACAGGCTCTTCGGGCCCATGGCAAGCACACGCTG 1176
QY 1168 ATGTCCCGACCTTCATAGATGAGACCGTGGCCAAACCCCTGGTACGCTGCACTGTGCC 1227
Db 1177 ATGGCAGCGCTGTGCTGCTCACTGAACAGACGCTGCCCCGTGGTCCCTGCGAGCGCATG 1236
QY 1228 ATCATCACGACTTCTGGACAGGGGACGCTGACTGCTCTCTGGACCAACCCAGCAAG 1287
Db 1237 TATCTCACAGAGCTTCTGGACGGGGGACGAGAGCTGTCTCTGATGCCCTGTGTGG 1296
QY 1288 CCCATCTCCCTGCCGAGGATCTGCCGGC-----GCCAGCTACACCTTGACCCAGCAG 1341
Db 1297 GCCCTGCCCTCCACAGGCTCCCGGGCGGATGSCCTGTACACAGCTGGACACGAG 1356
QY 1342 TGCAGAGTGGCTTTTGGCGTGGGTCTCAAGCCCTGTCTTACA-----TGCAGTAC 1392
Db 1357 TGCAGGAGAGATCTTTGGGCGCGGATTTCCGCACTGCCCCAACACCTGTGTGACGAGTC 1416
QY 1393 TGCACCAAGCTGTGGT--CACGGGAAGGCCAAGGACAGATGTTGTGCCAGACCCGC 1449
Db 1417 TGGCCCAAGCTTTTGGTCCACATGATGAGGGCTGAGCCCTGTGCCACAGAAAGATGGC 1476
QY 1450 CACTTCCCTGGCGGATGGACACAGCTGTGGCAGGGCAAGCTCTGCTTCAAGAGGGCC 1509
Db 1477 AGCTGCTGGCTGAGCGACACCGCTGCGGGCTGGGCACTCTGCTCAGAAGGACG 1536
QY 1510 TG-----CGTGGAGACACACCTCAACAGCACAGGTGGATGTTCTTGGGCCAA 1563
Db 1537 TGTCTACCTGAGGAGGAAGTGGAGAGGCCCAAGGCCCTGTAGATGAGGCTGGGCACCG 1596
QY 1564 TGGATCCCTATGCCCTGTGCGCACATGTGTGGGGCGGTGACGCTGGCCAGGAGG 1623
Db 1597 TGGGACCTGGGAGATGTTCTCGACCTGTGGAGGAGGATACAGTTTTCACACCGT 1656
QY 1624 CAGTGCACCAACCCACCTTGCACACGCGGGCAAGTACTGCGAGGAGGTGAGGGTGAAA 1683
Db 1657 GAGTGAAGGACCCCGACCTTCAGATGGAGAAATGATCTGCTGGGTGGAGGACCAAG 1716
QY 1684 TACCATCTCCTGATCTGGAGCCTGCCCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAG 1743
Db 1717 TACCAGTATGCCACAGGAGGAATGCC-----CCTACGGGAAGAGCTTCAGGGAG 1770
QY 1744 GAGCAGTGTAGGGCTTTCAACGGGTACAAACACAGCACACCGGCTTCACTCTCGCGGTG 1803
Db 1771 CAGCAGTGTGAGAGTATAATGCTACAATTTACACTGAGCAGGGAATCT--CCTG 1827
QY 1804 GCATGGGTGCCAAGTACTCTCGGCGTGTCTCCCGGGACAGAGTCAAGCTCATCTGCCGA 1863
Db 1828 CAGTGGGTCCCCAAGTATGCTGGGGTGTCCCGCGGGACCGCTGCAAGTGTCTTCCCGA 1887
QY 1864 GCCAATGSCACTGCTACTTCTATGTCTGTCACCCCAAGGTGGTGGACGACGCTGTGC 1923

Db 1888 GCCCGGGGAGGAGGAGTTCAAAAGTGTTCAGGCCCAAGGTGATGATGGCACCTGTGT 1947
QY 1924 TCTCTACTCTCCACTCGCTGTGTCCAGGCAAGTGCATCAAGGCTGGCTGTGATSGG 1983
Db 1948 GGGCCAGAAACACTGGCCATCTGTGCGTGGCCAGTGTGCAAGCGCGGTGTGATAT 2007
QY 1984 AACTTGGCTTCAAGAGAGATTCGACAAAGTGTGGGTGTGGGGGAGACATAAAGC 2043
Db 2008 GTGTGGTACTCGCTCGGAAGCTGACAAATGCGGGGTGTGTGGGGCAAGGCAATCC 2067
QY 2044 TGCAGAGAGTGTACTGGACTCTTCACCAAGCCCATGATGCTGCTCAATTTGCTGTGGCC 2103
Db 2068 TCCAGGAAGGTCTCGGGTCCCTCACCCCACTATATGCTACATGACATTTGTCAAC 2127
QY 2104 ATCCCGCAGCGGCTCAAGCATGACATCCGCGCAGCGCGGTTACAAAGGCTGTATCGG 2163
Db 2128 ATCCAGCTGTGGCACTAATATTGACGTGAAGCAGCGGAGCCACCGGGGTGTGCAAAC 2187
QY 2164 GATGACAACTACTGGCTCTGAAGAACAGCAAGCAGGAGTACCTGTCTCAAGGGCAATTC 2223
Db 2188 GATGGAACTACTGGCTGAAGACGCTGTATGGGCACTACCTGCTCAACGGCAACTG 2247
QY 2224 GTGTGTGCGGGTGGAGGAGCTGTGTTGAAGGCACTGTCTGCGGTACAGCGC 2283
Db 2248 GCATCTCTGCCATGACAGACATCTTGTGAAGGGAGCCATCTTGAATGACAGCGC 2307
QY 2284 AGGGCAGACGGGTGGAGAGCTTCAGAGCTTCCCGGCCCATCTCTGGAGCCCTGACCTG 2343
Db 2308 TCCATCGCCACCTGGAGCGCTGCAGAGCTTCGCGCCCTTGCAGAGCCCTTGACAGTG 2367
QY 2344 GAGGTCTCTCGT---GGGGAAGATGACAGCGCCCGGGTCCGCTACTCTTCTAICTG 2400
Db 2368 CAGCTCTCTGACAGTCCCTGGGAGGTCTTCCCCCAAAAGTCAATACACCTTCTTT 2427
QY 2401 CCAAAAGA 2408
Db 2428 CCTATGA 2435

RESULT 13
AAC90058
ID AAC90058 standard; DNA; 3008 BP.
XX AAC90058;
AC AAC90058;
DT 19-MAR-2001 (first entry)
XX Human METH2 coding sequence.
DE Human; METH2; metalloprotease; thrombospondin; angiogenesis inhibition;
KW cancer therapy; benign tumour; ocular angiogenic disease;
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Webber syndrome;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; fibromuscular dysplasia; wound granulation;
KW Crohn's disease; atherosclerosis; birth control; ss.
XX Homo sapiens.
XX WO200071577-A1.
PN 30-NOV-2000.
XX 25-MAY-2000; 2000WO-US14462.
PF 25-MAY-1999; 99US-0318208.
XX 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.

22-FEB-2000; 2000US-0183792.
 (HUMA-) HUMAN GENOME SCI INC.
 (SMIK-) SMITHKLINE BEECHAM CORP.
 (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 (IRUE-) IRELA-ARISPE L.
 (HAST-) HASTINGS G A.
 (RUBE-) RUBEN S M.
 (JONA-) JONAK Z L.
 (TRUL-) TRULLI S H.
 (FORN-) FORNWALD J A.
 (TERR-) TERRETT J A.

IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
 Fornwald JA, Terrett JA;
 WPI: 2001-025136/03.
 N-PSDB; AAC90058.

METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit angiogenesis in the treatment of disorders such as cancer, rheumatoid arthritis and psoriasis -
 Claim 14; Fig 2; 768pp; English.

The present sequence is the coding sequence of human METH2 (ME for metalloproteinase and TH for thrombospondin). METH2 can be used for inhibiting angiogenesis in an individual; and for treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis, vasculogenesis, granulations, adhesions, myocardial infarction, scleroderma, trachoma, vascular collaterals, arteriovenous malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, fibromuscular dysplasia, wound granulation, Crohn's disease or atherosclerosis. METH2 can also be used in birth control. METH2 can also be used in diagnostic methods for the prognosis of cancer.

Sequence 3008 BP: 617 A; 924 C; 948 G; 514 T; 5 other;

Query Match 23.3%; Score 664.4; DB 22; Length 3008;
 Best Local Similarity 59.2%; Pred. No. 3.4e-129;
 Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;

QY 157 CTCATTTTCAGATCACAGCATTTTCAGAGGACTTTTACCTTACACTGACGCGCGATGCT 216
 DB 160 CTGGCGTCCACCTGTGCGCTTCGGCAAGGGCTTCGTGTGCGCTTGGCGCCGACGAC 219
 QY 217 CAGTTCTTGGCTCCGCGCTTCCACTGAGCATCTGGCGGTCCTCCAGGGGCTCACC 276
 DB 220 AGCTTCTTGGCGCGCGAGTTCAAGATGAGCGGCTCGGGGCTCCGGCGGGGCG-ACC 276
 QY 277 GGGGCTTTCAGACCTTCGACGCTTCCAGTACGAGCATCTGGCGGTCCTCCAGGGGCTCACC 336
 DB 277 GGGGCGAGCGGGGCTGCGCGCTGTCTTTTTCGGGACCTGTAATGGAGCGCGAG 336
 QY 337 TCGTTCGCTGTGAGCCTTGGCGGGGCTCCGCGAGCCTTTGGCTACCGAGGGGCG 396
 DB 337 TCGTTCGCGGGGTGAGCCTGTGCGGGCTGAGCGGCTTCTTCTGCTGGAGCGGCGAG 396
 QY 397 GAGTATGTCATTCGCGCTGCCAATGCTAGCGCGCGGGGCGGCGAGCGCAACAGCGAG 456
 DB 397 GAGTATACCATTCACGCGCGAGGGGCGGGGGGCTTCCCTGCTCAGCGCGACCGGCTGAG 456
 QY 457 GCGCACACCTTCTCCAGCGCGGGGTGTTCCGGCGGGGCTTCCGGAGACCCACCTCT 516
 DB 457 CGCTGGGGTCCCGCGAGCGCGCGCTTCCCGGAGGACCGAGTGGGAGGTGGAGAGC 516
 QY 517 CGCTCGGGGTG-CCCTCGGGCTGGAAACCGCGCATCTACGGGCGCTGGACCTTACAA 575
 DB 517 GGAGAGGCTCAGAGGCGAGAGAGAGAGACACCGAGGAGACAGCGGGAGAGAGACCAA 576

576 GCCCGCGGGCGGGCTTCGGGGAGAGATCTAGCCGCGCGCAGG-----TCTGGGCGCGCC 630
 DB 577 GAAGAGGAGGAGAGAGGCGCTAGCGAGCGCCACCGCCCTTGGGGGCGCACGAGTAGGACC 636
 QY 631 AAGCGTTTCGTGTCTATCCCGCGTACGTGAGAGCGTGTGTGGTGGCGGAGAGTCAATG 690
 DB 637 AAGCGTTTCGTGTGCTGAGGCGGCTTCTGAGAGCGTGTGTGGCGGAGTGGCTCATG 696
 QY 691 GTCAGTTTCCACGCGCGGACCTGGAATATATCTGCTGACGCTGTGTCGCAAGCGGCGG 750
 DB 697 GCTGCTTCTACGGGCGCGACCTGACAGACCATCTGACGTTAATGTCTGTGGAGCG 756
 QY 751 CGACTCTACGCGCATCCAGCATCTCAACCCATCAACATCGTTGTGGTCAAGGTGCTG 810
 DB 757 CGAATCTACAAGCACCCCGCATCAAGAATTCATCAACCTGATGGTGGTAAAGTCTG 816
 QY 811 CTTCTTAGAGATCGTGACTCCGGGCGCAAGTCCAGCGCAATCGCGCCTGACGCTGCGC 870
 DB 817 ATCGTAGAAGATGAATAATGGGCGCGAGAGGTGTCCGACAAATGGGGCTTTACACTCG 876
 QY 871 AACTTCGTGCTGGCGAGAGAGCTGAACAAAGTGAAGTGAACAGCAACCCGAGTACTG 930
 DB 877 AACTTCTGCAACTGGCAGCGCGCTTTCAACCGAGCGACCGCCACCCAGAGCACTAC 936
 QY 931 GACACTGCCATCTCTTCACGAGGAGGACTGTGTGG-----AGCCACCACTGTGACAC 987
 DB 937 GACAGCGCATCTCTTCACGAGAGAGTCTGTGGCGAGGAGGGCTGTGTGACACC 996
 QY 988 CTGGGATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGTCTGTCTATTGAG 1047
 DB 997 CTGGGTGTGACAGATCGGGACCAATTTGTGACCCCAACAAAGCTGTCTGATCGAG 1056
 QY 1048 GAGATGGGCTTCCATCAGCTTACACTGCGGCGAGCTGGGCGACGTTTCAACATG 1107
 DB 1057 GATGAGGGGCTCCAGCGCGCCACACCTGGCCCATGAATAGGAGCAGCTCTCAGCATG 1116
 QY 1108 CCCATGACAAATGAAAGTCTGTGAGGAGTGTGTTGGAGGCTCCGAGCCACCAACATG 1167
 DB 1117 CCCAGCAGGACTCCAAAGCCCTGCACAGGCTCTTCGGGCGCATGGGCAAGCAACAGTG 1176
 QY 1168 ATGTCGCGGACCTCATCCAGATCGACCGTGCACCCCTGTGTGACCTGTGATGTGCG 1227
 DB 1177 ATGGCACCGCTGTGCTCCACCTGAACCAAGACGCTGCTGCTGCGCGGCGCATG 1236
 QY 1228 ATCATCAGCAGCTTCTGACAGCGGCGAGTGTGCTCTGTCGACCAACCCAGCAAG 1287
 DB 1237 TATCTCAGAGAGTCTTGGACGCGGCGAGAGACTGTCTCTGGATGCCCTGTGCG 1296
 QY 1288 CCATCTCCCTGCGCGGAGTCTGCGGGG-----GCCAGCTACACCTGAGCCAGCAG 1341
 DB 1297 GCGCTGCCCTCCCAAGCGCTCCCGGGCGCATGGCCCTGTACAGCTGACCAAGCAG 1356
 QY 1342 TGGAGCTGCTTTTGGCTGGGCTCCAGCCCTGTCCTTACA-----TGCAAGTAC 1392
 DB 1357 TGCAGCAGATCTTTGGCGGATTTCCGCCACTGCCCCAACACCTCTCTCAGGAGCTC 1416
 QY 1393 TGCACCAAGCTGTGG-----CACCGGAGGCGCAAGGACAGATGCTGTGCGAGACCCG 1449
 DB 1417 TGGCGCGAGTGTGGTGGCAGACTGATGGGCTGAGCCCTGTGCGACAGAGATGCG 1476
 QY 1450 CACTTCCCTGGGCGGATGGCACCAAGCTGTGGCGAGGCGCAAGCTGTGCTCAAGGGCG 1509
 DB 1477 AGCTGCCCTGGCTGACGCGCGCGCTGCGGGCTGGGCGCTGTGCTCAAGAGCAGC 1536
 QY 1510 TG-----CGTGAGAGACACACCTCAACAGCAGAGGTTGGTTCCTGGGCGCAAA 1563
 DB 1537 TGTCTACCTGAGGAGGAGTGGAGGCGCCAAAGCCCGTGGTATGAGGAGTGGGCGCG 1596
 QY 1564 TGGGATCCCTATGGCCCTGCTGCGGACATGTGTGGGGGCGTGCAGCTGCCAGGAGG 1623
 DB 1597 TGGGAGCCTGGGAGAGATGTTCTCGGACCTGTGGAGGAGGAGTACAGTTTTCACACCG 1656
 QY 1624 CAGTGACCAACCCACCCTGCCAACGGGGGCAAGTACTGCGAGGAGTGTAGGCTGAAA 1683

QY 991 GGCATGGCTGATGTGGTACCATGTGTGACCCCAAGAGAGCTGCTCTGTCATTGAGGAC 1050
Db 1503 GGAATGGCAGATGTTGGAAACCGTATGTGACCCCAAGAGAGCTGCTCTGTCATTGAGGAC 1562
QY 1051 GATGGCTTCCATCATGACCCCTTCCACCACTGCCACAGAGTGGGCCACAGTGTTCACATGCCC 1110
Db 1563 GATGTTGGCAAGCGCCCTTCCACCACTGCCACAGAGTGGGCCACAGTGTTCACATGCCC 1622
QY 1111 CATGACATGTGAAGTCTGTGAGGAGTGTGGGAGCTTCCAGCCCAACCATGATG 1170
Db 1623 CACGATGATGTGAAGTGTGGCAGCTGTGAATGGTGTGAGTGGCGATTCATCTGATG 1682
QY 1171 TCCCGGACCCCTCATCATGACATGACCGTGCACACCCCTGTCAGCCTCAGTGTGGCCATC 1230
Db 1683 GCCTCGATCTCTCCAGCTTAGACCATAGCAGCCCTGTCACCTTGCAGTGGCTACATG 1742
QY 1231 ATCAAGGATCTCTGGACAGGGGACCGGTGACTGCTCTGGACCAACCCAGCAAGCC 1290
Db 1743 GTCAGTCTCTTCTAGATAATGGACAGGGGAATGTTGATGGACAAAGCCCAAGATCCA 1802
QY 1291 ATCTCCTCTCCGAGGATCTGCGGGGCCAGCTAGACCTGAGCCAGCACTGCGAGCTG 1350
Db 1803 ATCAAGTCTCTCTGATCTTCCCGTACCTTGTACGATGCCAAGCCGCTGTCAGTTT 1862
QY 1351 GCTTTTGGCGTGGGTCCAAAGCCCTGCTCTTACATGCAAGTAC ---TGACCAAGCTGTGG 1407
Db 1863 ACATTCGGAGAGGAATCCAAAGCACTGCGCTGATGCAAGCAGCAGTGTACCTGTGG 1922
QY 1408 TGCACCGGAAGGCCAAGGACAGATGTTGCCAGACCCGCCACTTCCCTGGCCGAT 1467
Db 1923 TGCACCTGGCACTCTGGTGGCTTACTGGTGGTCCCAAAACAAACACTTCCCTGGCCAGAT 1982
QY 1468 GGCACGAGCTGTGGGAGGGCAAGCTCTCCCTCAAAGGGGCTGCTGGTGAGAGACACAAC 1527
Db 1983 GGCACGAGCTGTGGAGAGGAAGTGTGTGCTGCTGCAAGTGGTGGTGAACAGACAGAC 2042
QY 1528 CTCACACAGCACAG-----GCTGATGTTCTCTGGGCCAAATGGGATCCCTATGGCCGC 1581
Db 2043 ATGAACATTTTGTCTACTCTCTTCTGTAAGCTGGGAGCCATGGGACCTGGGGAGAC 2102
QY 1582 TGTCTCGGCACATGTGTGGGGGCTGTCAGCTGGCGAGGAGCTGACCAACCCCAAC 1641
Db 2103 TGTCTAAGAACCTGTGGTGGTGGAGTTCATACACAATGAGAGATGTGACAACCCAGTC 2162
QY 1642 CTGTCACAGGGGGCAAGTACTGCGAGGAGTGTAGGGTCAATACCGATCTGCAATCTG 1701
Db 2163 CCAAGAAGCGGGAAGTACTGTGAAGGCAAAAGCTGCGTACAGCTCTGTAACATC 2222
QY 1702 GAGCCCTGCCAGCTCAGCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGAGCTTTC 1761
Db 2223 GAGGACTGTCCAGACAATAA---CGGAAACGTTTCAGAGGAGCAGTGGCGGCGAC 2279
QY 1762 AAGCGCTACAAACACAGCACCAACCGGCTCACTCTCGCGGTGGCATGGTGGCCAAAGTAC 1821
Db 2280 AATGAGTTTTCACAACTTCTTCTGGATGAGCCCACTGTAGTGGACACCCCAAGTAC 2339
QY 1822 TCCGGCGTGTCTCCCGGACAAAGTCAAGCTCATCTGCGGAGCCAAATGCGACCTGGCTAC 1881
Db 2340 GCGGGGTCTCGCAAGGACAGGTGCAAGCTCAGCTGTGAAGGCAAAAGGCAATTTGGCTAC 2399
QY 1882 TTTATGTGTGCGACCCCAAGTGTGGAGCGCAGCTGTCTCTCTGACTCCACCTCC 1941
Db 2400 TTTTCTGTCTTACAGCCCAAGGTTGTAGATGGCACTCCCTGTAGTCCAGACTCTACCTCT 2459
QY 1942 GTCTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGTGATGGGAACCTGGGCTCCCAAGAG 2001
Db 2460 GTCTGTGTCAAGGCGAGTGTGAAAGCTGTGTGATGGCATCATAGACTCCCAAGAG 2519
QY 2002 AGATTGCAAGTGTGGGTGTGTGGGGAGACATATAGAGCTGCAAGAGTGTACTGGA 2061
Db 2520 AAGTTTGTATGTTGGCTTTTGTGGAGGAACGGTTCCACATGCAAGAGATGTGAGGA 2579

QY 2062 CTCCTTCAACAAGCCCATGCATGGCTACAATTTCTGGTGGCCATCCCGGAGGGCTCA 2121
Db 2580 ATAGTCACTAGTACAAGACCTGGGTATCATGACATTTGTACAATTTCTGCTGGAGCCACC 2639
QY 2122 AGCATGACATCCGCCAGCGGGTTACAAGGGGTGATCGGGGATGACACTACTGGCT 2181
Db 2640 AACATTTGAAGTGAACATCGGAATCAAAGGGGGTCCAGAAACAATGGCAGCTTCTGGCT 2699
QY 2182 CTGAAGAAGCAGCAAGGCAAGTACCTGCTCAAGGGCATTTCTGGTGTCTGGCGGTGGAG 2241
Db 2700 ATTAGAGCGCTGATGGTACCTATATTTCTGAATGGAACCTTCACTCTGTCCACACTAGAG 2759
QY 2242 CGGACCTGGTGGTGAAGGAGTCTGCTGGGTACAGCGGACGGGACAGCGGTGGAG 2301
Db 2760 CAAGACCTCACCTACAAAGTACTGCTTAAGGTACAGTGGTCTCTGGGTGGCTGGA 2819
QY 2302 ASCCTGCAAGCTTCCCGGCCCATCTCGGAGCGCTGAGGAGTCTCTCGGTGGG 2361
Db 2820 AGAATCCGAGCTTTAGTCCACTCAAGAACCTTACCATCCAGTCTTATGGTAGGC 2879
QY 2362 AAGATGACACACCCCGGCTCGCTACTCTTCTTATCTGCCCAAGACCTCGGGAGGAC 2421
Db 2880 CATGCTCTCGGACCCCAAAATTAATTTACCTACTTATGAAGAAGACAGAGTCACTC 2939
QY 2422 AAGTCTCTC 2431
Db 2940 AAGGCCATTC 2949

RESULT 15
AAC90067
ID AAC90067 standard; DNA; 4180 BP.
XX AAC90067;
XX
XX
XX 19-MAR-2001 (first entry)
XX
XX D67076 cDNA clone.
XX
XX
KW METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW cancer therapy; benign tumour; ocular angiogenic disease;
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW plaque neovascularisation; telangiectasia; haemophilic joint; ESP;
KW angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW Crohn's disease; atherosclerosis; birth control; ss.
XX Unidentified.
XX
XX WO200071577-A1.
XX
XX 30-NOV-2000.
XX
XX 25-MAY-2000; 2000WO-US14462.
XX
XX 25-MAY-1999; 99US-0318208.
PR 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.
PR 22-FEB-2000; 2000US-0183792.
XX
(HUMA-) HUMAN GENOME SCI INC.
PA (SMK) SMITHKLINE BEECHAM CORP.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA (IRUE/) IRUELA-ARTISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
PA (JONA/) JONAK Z L.
PA (TRUL/) TRULLI S H.

PA (FORN/) FORNWALD J A.
PA (TERK/) TERRETT J A.
XX Iruela-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
PI Fornwald JA, Terrett JA;
XX WPI; 2001-025136/03.
XX METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT inhibit angiogenesis in the treatment of disorders such as cancer,
PT rheumatoid arthritis and psoriasis -
XX Claim 7; Pages 543-546; 768pp; English.
XX The present invention relates to human METH1 and METH2. (ME for
CC metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC The present sequence is an expressed sequence tag (EST) for METH. METH
CC can be used for inhibiting angiogenesis in an individual, and for
CC treating cancer, benign tumours, an ocular angiogenic disease,
CC rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC scleroderma, trachoma, vascular adhesions; myocardial angiogenesis;
CC coronary collaterals, cerebral collaterals; arteriovenous malformations,
CC ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
CC fibromuscular dysplasia, wound granulation, Crohn's disease or
CC atherosclerosis. METH can also be used in birth control. METH can also
CC be used in diagnostic methods for the prognosis of cancer.
XX Sequence 4180 BP; 1078 A; 1051 C; 1146 G; 905 T; 0 other;
Query Match 22.2%; Score 634.4; DB 22; Length 4180;
Best Local Similarity 60.6%; Pred. No. 6.7e-123;
Matches 1097; Conservative 0; Mismatches 701; Indels 12; Gaps 3;
QY 631 AAGCGTTTCGTATCCCGGCTAGGTGGAGACGGTGTGGTGGGAGACGAGTCAATG 690
DB 1143 AAGCGATTGTGTCCAGCCCGCTTATGTGGNAACCATGCTCGTAGTGCACCATG 1202
QY 691 GTCAAGTTTCCAGCGCGGACCTGGAAACATTATCTGTGACGCTGTGGCAACGGGGCG 750
DB 1203 GCCACTTCCAGCGCAGCGGTCTAAGCAATACCTCTAACCTGTCTCGGTGGCAGCC 1262
QY 751 CGACTACCGCCATCCAGACATCCTCAACCCCATCAACATCGTTGGTGTCAAGTGTG 810
DB 1263 AGGTTTACAAGCATCCAGCATAGGAATTAATAGCCCTGGTGGTGTGAAGATCTTG 1322
QY 811 CTTCTTAGATCTGTACTCCGGCCCAAGTCTACCGCAATGGCCCTGACGCTGCCG 870
DB 1323 GTCATATACGAGGACGAGAGGACCAAGAGTTACCTCCATGACGCTCTCACCTTCGG 1382
QY 871 AACTTCTGTGCTGGCAGAGAACTGAACAAAGTGAAGTCAACAGCACCCCGAGTACTGG 930
DB 1383 AATTCTGACGCTGGCAGAACACACACAGCCCAAGTCAACCGGATCCAGAGCACTAT 1442
QY 931 GACACTGCCATCTCTTCACAGCAGGACCTGTGTGGAGCCACCACTGTGACACCCCTG 990
DB 1443 GACATGCAATCTGTTCACAGACAGAGTTATGTGGCTCCCAACACAGCTGTGACACTC 1502
QY 991 GGCATGGCTGATGTGGTACCATCTGTGACCCCAAGAGAGCTCTCTGTCATTGAGGAC 1050
DB 1503 GGAATGGCAGATGTGTGAACCGTATGTGACCCCAAGAGAGGAGTCTCTGCTCATAGAGAT 1562
QY 1051 GATGGGCTTCCATCAGCTTCCACACTGCCACAGCTGCCAGGCTGGGCCACGTGTTCACATGCC 1110
DB 1563 GATGGTTTGAAGCGCTTCCACACAGCCCATGAATTGGGCCATGTGTTAATCATCCG 1622
QY 1111 CATGACATGTGAAGTCTGTGAGAGGTGTTTGGGAGCTCCAGACCCACACCATCATG 1170
DB 1623 CACGATGATGTGAAGCACTGTGCCAGCTTGAATGTGTGATGTGGCGATTCTCTCTGATG 1682
QY 1171 TCCCGACCCCTCATCCAGATCGACCGTGCACCAACCCCTGGTTCAGCTGCAGTGTGCCATC 1230

DB 1683 GCCTCGATGCTCTCCAGCTTAGACATAGCAGCCCTGGTGCACCTTGCAGTGCCTACATG 1742
QY 1231 ATCACCAGACTTCTCTGGACAGCGGCACGGTGCACCTCTCTGGACCAACCCAGCAAGCCC 1290
DB 1743 GTCAGCTCTCTCTAGATATATGACACAGGGGAATGTTTGTGACACAGCCCAAGATCCA 1802
QY 1291 ATCTCCCTGCCCCGAGGATCTGCCGGGCGCCAGCTTACACCTTGAGCCAGAGTGGGAGCTG 1350
DB 1803 ATCAAGCTCCCTCTCTGATCTTCCCGGTACCTTGTACGATGCCAAACCCGCACTGTGAGTT 1862
QY 1351 GCTTTTGGCTGGCTCCAAAGCCCTGCTTACATGACAGTAC---TGCACCAAGCTCTGG 1407
DB 1863 ACATTCGGAGAGGAATCAAGCACTGCCCTGATCCAGCAGCACATGTACTACCTCTGG 1922
QY 1408 TGCACCGGGAAGGCCAAGGGACAGATGTTGTCAGACCCCGCCACTTCCCTGGGCGCAT 1467
DB 1923 TGCAGTGCACCTCCGGTGGCTTACTGTGTGCCAAACAAACACTTCCCTTGGCAGAT 1982
QY 1468 GGCACCACTGTGGGAGGCAAGCTCTGCTCTAAAGGGGCTCGCTGGAGAGACACAAAC 1527
DB 1983 GGCACCACTGTGGAGAGGAAGTGTGTCTAGTGGCAAGTGGTGAACAGACAGAC 2042
QY 1528 CTCACAAAGCACAG-----GTTGATGTTCTCTGGGCCAAATGGATCCCTATGGCCCC 1581
DB 2043 ATGAAGCAATTTTGTACTCTCTGTTTCATGGAAGCTGGGGACCATGGGACCGTGGGAGAC 2102
QY 1582 TGCTCGGCACATGTGTGGGGCTGTCAGCTGGCAGAGGAGGAGTGCACCAACCCCAACC 1641
DB 2103 TGCTCAAGAACCTGTGTGGTGGATTCATATACAAATGAGAGAATGTGACAAACCCAGTC 2162
QY 1642 CCGTCCACAGGGGCAAGTACTGCGAGGAGTGGGTGAATACCGATCCCTGCAATCTG 1701
DB 2163 CCAAAAGAACGGGGAAGTACTGTGAAGGCAACAGAGTCCGCTACAGGTCTCTGTAACATC 2222
QY 1702 GAGCCCTCCCGAGCTCAGCCTCCGGAAGAGCTTCCGGGAGGAGCAGTGTGAGGCTTTC 1761
DB 2223 GAGGACTGTCCAGACAATAA---CGGAAAACGTTTCAGAGAGGAGCAGTGTGAGGCGCAC 2279
QY 1762 AACGGCTACAAGCACAGCACCAACCGGCTCAGTCTGCGCGTGGCATGGTGGTCCCAAGTAC 1821
DB 2280 AATGAGTTTTCCAAAGCTTCTCTTTGGGAATGAGCCCACTGTAGAGTGACACCCCAAGTAC 2339
QY 1822 TCCGGCTGTCTCCCGGACAAAGTCAAGTCACTCTCCGAGCCAAATGGCACTGGCTAC 1881
DB 2340 GCCGGCTCTCCCAAGAGACAGTGTCAAGTCACTCTGTGAAGCCAAAGGATTTGGCTAC 2399
QY 1882 TTCTATGTCTGGCCCAAGGTGTGGACCGGACGCTGTCTCTGCTGATCTCCACCTCC 1941
DB 2400 TTTTCTGCTTACAGCCCAAGGTTCTAGATGGCACTCCCTGTAGTCCAGACTCTACCTCT 2459
QY 1942 GTCTGTGCCAAGGCAAGTGCATCAAGCTGGCTGTGTGATGGGAACCTGGGCTCCGAAGAG 2001
DB 2460 GTCTGTGCCAAGGCAAGTGTGAAAGCTGGCTGTGTGATCGCATATAGACTTCCAAAAG 2519
QY 2002 AGATTGCAAGCTGTGGGTGTGTGGGGAGACAAATAAGAGCTGCAAGAAAGTGAAGTGA 2061
DB 2520 AAGTTGATAGTGTGGCTTGTGGAGGAACGCTTCCACATGCAAGAGATGTGAGAG 2579
QY 2062 CTCCTTCAACAAAGCCATGCATGGCTTACAATTTCTGTGGTGGCCATCCCGAGGCGCTCA 2121
DB 2580 ATAGTCACTAGTACAAGACCTGGGTATCATGACATTTGTACAAATTTCTCTCTGGAGCCAC 2639
QY 2122 AGCATGCATCCGCCAGCGGTTTACAAGGGCTGTACAAAGGGTGTACGGGATGACAACCTACCTGCT 2181
DB 2640 AACATTGAAGTGAACATCGGAATCAAGGGGGTCCAGAAACATGGGACGTTTCTGGCT 2699
QY 2182 CTGAAGAACAGCAAGCAAGTACCTGTCTCAACGGGCAATTTCTGGTGTGGCGGTGAG 2241
DB 2700 ATTAGAGCGCTGTGTGTTACCTATATTCTGAATGGAACCTTCACTCTGTCCACACTAGAG 2759
QY 2242 CGGAGCTGTGTGTAAGGGCAGTCTGTGCGGTACAGCGGCACGGGCACACCGTGTGAG 2301
DB 2760 CAAGACCTCACCAAGAGTACTGTCTTAAGGTACAGTGGTTCCTCGGCTCGCTGGAA 2819

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 16:17:29 ; Search time 7234 Seconds
(without alignments)
11477.786 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtctctgtggcctact.....gcctcttgagccgtgtctga 2853

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rtd.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2851.4	99.9	2853	9	HS315733
2	2817.6	98.8	2930	6	AX342635
3	2817.6	98.8	2937	6	AX319860
4	960.4	33.7	157963	9	AP002986
5	960.4	33.7	172905	2	AC025130
6	896.2	31.4	170682	2	AC023429
7	783	27.4	182656	2	AC101990
8	779	27.3	178764	2	AC126507
9	777.2	27.2	28000	9	AP003459
10	773.6	27.1	3927	10	BC009667
11	664.4	23.3	2670	6	E5282
12	664.4	23.3	2670	6	E58655
13	664.4	23.3	3711	9	AF060153
14	634.4	22.2	4180	10	D67076
15	629.8	22.1	4659	9	AF207664
16	628.2	22.0	3430	9	AF060152
17	628.2	22.0	4014	6	E29406
18	628.2	22.0	4447	9	AF170084
19	627.2	22.0	2184	6	E29666
20	626.6	22.0	2853	6	E58656
21	626.6	22.0	4670	9	BC036515
22	625	21.9	4309	9	AB037767
23	612.8	21.5	4878	10	AF149118
24	606.4	21.3	2857	10	AF304446
25	597.4	20.9	178764	2	AC126507
26	592	20.8	3638	10	AF175282
27	583.2	20.4	4193	9	AF148213
28	583.2	20.4	4301	9	AB014588
29	581.6	20.4	4407	6	AX376284
30	562.6	19.7	3321	10	BC027773
31	549	19.2	3002	10	AF140673
32	549	19.2	5533	9	AF142099
33	417.6	14.6	5139	9	AB037733
34	417.6	14.6	5808	6	AX319857
35	416	14.6	3674	9	AF261918
36	402.4	14.1	1251	10	AB042271
37	382.4	13.4	2838	9	HS0802617
38	375.6	13.2	5724	6	AX392458
39	375.6	13.2	6060	6	AX392460
40	374	13.1	172905	2	AC025130
41	367.8	12.9	1560	9	AF141293
42	322	11.3	2554	9	AK094057
43	319.6	11.2	1365	10	AB042272
44	316	11.1	3308	6	AX480930
45	316	11.1	4518	6	AX319858

ALIGNMENTS

RESULT 1	HS315733	2853 bp	mRNA	linear	PRI 01-MAR-2002
LOCUS	HS315733				
DEFINITION	Homo sapiens mRNA for metalloprotease disintegrin 15 (ADAMTS15 gene).				
ACCESSION	AJ315733				
VERSION	AJ315733.1	GI:19171175			
KEYWORDS	ADAMTS15 gene; disintegrin; metalloprotease; thrombospondin.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Cal.S., Obaya,A.J., Llanazares,M., Garabaya,C., Quesada,V. and Lopez-Otin,C.				

TITLE Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains
JOURNAL Gene 283 (1-2), 49-62 (2002)
MEDLINE 21856482
PUBMED 11867212
REFERENCE 2 (bases 1 to 2853)
AUTHORS Cal.S.
TITLE Direct Submission
JOURNAL Submitted (26-Jun-2001) Cal S., Biochemistry and Molecular Biology, University of Oviedo, Campus del Cristo, Asturias. 33006, SPAIN
LOCATION/Qualifiers
FEATURES
source
1..2853
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1..2853
/gene="ADAMTS15"
CDS
1..2853
/gene="ADAMTS15"
/codon_start=1
/product="metalloprotease disintegrin 15 with thrombospondin domains"
/protein_id="CAC86014.1"
/db_xref="GI:19171176"
/translation="MILLGLTLAFAGRTAGSEPEREVYVVRIRLDPDINGRRYYWRG
PEDSGQGLIFQITAFQEDFYHLHTPAQFLAPAEVSTHGLVPLQGLTGGSDLRRCF
YSGDVNAEPDSFAVSLQGLRGAFYRGAEVVISPLPNASAPAOAQRNSOGAHLQRR
YGPVSGDPTSGKSVAGSNPALLRALDYPKRRAGFGESESRRSRRGRKRFVSIPIR
YVETLVADSESVKFGADLEHYLLTLATAARLYRHPSILNPINIVVKKLLLRD
SGPKVTNALTLRNECAWOKLNKVRDPEWTDAILFTQDLCGATCTDLGMAD
VTGMDPKRSVSIEDDGLPSFTTAHGLHGFNPMHDNVKVEEVEFGKLRNHHMSP
TLQIDRANPWSACSAIIITDFLSDQGLDQPSKIPSLDEPLDIPGASYTLSSQCEL
AFVGSKPCPYMOCYKLMCTKAKGMVQOTRHFPAWDTSCGEGKCLKGACVERH
NLNKRVDGSWAKMDYGPCSRTGGVOLARRQCTNPTPANGVKYCEVRVYKRSN
LEPCPSASGKSPREDCAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCKLICRANG
TGYPYVLAKPVVDCTLCSPDSTSCVQGCIRAGCDNGLSKRPDKCGVCGDNKSC
KRYTGLTFPMHGNFVAPAGASSIDIRQRYKGLIGDNLALKNQKQKLLNGH
FVYSAYERDLVWKSLLRYSGLTAVESLOASRPLEPLETVESVGRNMTPRVRYSF
YLPKREDKSSHPKDPGSPVSLHNSVLSNOVEQDPRPPARVWAGSWGPCSASCG
SGLOKRAVDCGSGAGTGPACDAHRPVEVETACGPECTWELSAWSPCKSGRGFQ
RRSLKCVGHGRLLRDQCNLHRKPOELDFCVLRPC"
BASE COUNT 522 a 921 c 903 g 507 t
ORIGIN
Query Match 99.9%; Score 2851.4; DB 9; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTTCTGCTGGGCATCCTTAACCTTGGCTTCGCCGGGCGAACCGCTGGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGCATCCTTAACCTTGGCTTCGCCGGGCGAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTCCATCCGACTGGACCGCGGACATTAACGCGCGCGCTAC 120
DB 61 CCAGAGCGGAGGTAGTCTCCATCCGACTGGACCGCGGACATTAACGCGCGCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCACAGCATTT 180
DB 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTTCAGATCACAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACTCAGCCGGATCCTCAGTCTTGGCTCCCGCTCTCC 240
DB 181 CAGGAGGACTTTTACCTACACTCAGCCGGATCCTCAGTCTTGGCTCCCGCTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCGTCCAGGGGCTACCGGGGCTCTCTTCAGACCTGCGACGC 300
DB 241 ACTGAGCATCTGGGCGTCCCGTCCAGGGGCTACCGGGGCTCTTCAGACCTGCGACGC 300
QY 301 TCGTTCTATTCTGGGAGCTGAACCGCGAGCCGACTCGTTCGCTGCTGTGAGCCTGTGC 360
DB 301 TCGTTCTATTCTGGGAGCTGAACCGCGAGCCGACTCGTTCGCTGCTGTGAGCCTGTGC 360
QY 361 GGGGGGCTCCCGGAGCCTTTGGGTACCGAGGCGCGAGTATGTCAATAGCCCGCTGCC 420

DB 361 GGGGGGCTCCCGGAGCCTTTGGTACCGAGCGCGAGTAGTATGTCAATAGCCCGCTGCC 420
QY 421 AATGTAGCGCGCGCGCGCGCGAGCAACAGCAGGCGCGACACCTTCTCCAGCGCGG 480
DB 421 AATGTAGCGCGCGCGCGCGCGAGCAACAGCAGGCGCGACACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGCGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGG 540
DB 481 GGTGTTCCGGCGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGG 540
QY 541 AACCCCGCATCTACGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGG 600
DB 541 AACCCCGCATCTACGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTCGGGG 600
QY 601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCAAGGTTTCTGTCTATCCCGCGTACGTG 660
DB 601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCAAGGTTTCTGTCTATCCCGCGTACGTG 660
QY 661 GAGACGCTGGTGTGTCGCGGACAGTCAATGTGTCAGTTTCCACGGCGCGAGCTTGAACAT 720
DB 661 GAGACGCTGGTGTGTCGCGGACAGTCAATGTGTCAGTTTCCACGGCGCGAGCTTGAACAT 720
QY 721 TATCTGTGACGCTGCTGGCAACGCGCGGACTCTACCGCATCCAGCATCTCAAC 780
DB 721 TATCTGTGACGCTGCTGGCAACGCGCGGACTCTACCGCATCCAGCATCTCAAC 780
QY 781 CCATCAACATCGTTGTGTTGTTCAAGTGTCTTCTTAGAGATCGTGTCTCCGGGCGCAAG 840
DB 781 CCATCAACATCGTTGTGTTGTTCAAGTGTCTTCTTAGAGATCGTGTCTCCGGGCGCAAG 840
QY 841 GTACCGGCAATGCGGCGCTGACGCTGGCAACCTTCTGTGCTGGCAAGAAAGCTGAAC 900
DB 841 GTACCGGCAATGCGGCGCTGACGCTGGCAACCTTCTGTGCTGGCAAGAAAGCTGAAC 900
QY 901 AAGTGTGAGCAACACACCCGAGTACTGGGACACTGCATCTCTTCCACAGCAGGAC 960
DB 901 AAGTGTGAGCAACACACCCGAGTACTGGGACACTGCATCTCTTCCACAGCAGGAC 960
QY 961 CTGTGTGGAGCCACCCTGTGACACCTTGGCATGGTGTGTGGTGTGGTGTGGTGTGGTGTGG 1020
DB 961 CTGTGTGGAGCCACCCTGTGACACCTTGGCATGGTGTGTGGTGTGGTGTGGTGTGGTGTGG 1020
QY 1021 CCCAAGAGAGCTGCTCTGCTATTGAGGAGATGGCTTCCATCAGCTTCCACACTGCC 1080
DB 1021 CCCAAGAGAGCTGCTCTGCTATTGAGGAGATGGCTTCCATCAGCTTCCACACTGCC 1080
QY 1081 CACGAGCTGGGCGACGTGTTCAACATGCCCCATGACAAATGTGAAAGTCTGTGAGGAGTG 1140
DB 1081 CACGAGCTGGGCGACGTGTTCAACATGCCCCATGACAAATGTGAAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGAGAGCTCCGAGCCAAACACATGATGTCGCCGACCTCATCCAGATCGACCTGCC 1200
DB 1141 TTTGGAGAGCTCCGAGCCAAACACATGATGTCGCCGACCTCATCCAGATCGACCTGCC 1200
QY 1201 AACCCCTGTGAGCTGCGAGTCTGCTCATCATCACGACTTCTTGACAGCGGCGACGCT 1260
DB 1201 AACCCCTGTGAGCTGCGAGTCTGCTCATCATCACGACTTCTTGACAGCGGCGACGCT 1260
QY 1261 GACTGCTCTTGAGCAACCCAGCAAGCCATCTCCCTGCCCGAGATCTGCCGGGCGCC 1320
DB 1261 GACTGCTCTTGAGCAACCCAGCAAGCCATCTCCCTGCCCGAGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCTGAGCCAGCAGTGGGCTGCTTTGGGCTGGGCTTCCAAAGCTTGTCT 1380
DB 1321 AGCTACACCTGAGCCAGCAGTGGGCTGCTTTGGGCTGGGCTTCCAAAGCTTGTCT 1380
QY 1381 TACATCAGTACTGTACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGTTGTGC 1440
DB 1381 TACATCAGTACTGTACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGTTGTGC 1440
QY 1441 CAGACCGGCACTTCCCTTGGGCGGATGACACAGCTGTGGGCGGAGGCAAGCTTGCCTC 1500
DB 1441 CAGACCGGCACTTCCCTTGGGCGGATGACACAGCTGTGGGCGGAGGCAAGCTTGCCTC 1500

Db 1441 CAGACCCGCCACTTCCCTGGGCCGATGGCACCAGCTGTGGCGAGGCAAGCTCTGCCTC 1500
 QY 1501 AAAGGGCTCGTGGAGAGACAACTCAACAGCAGGAGGATGCTTCTGGGC 1560
 Db 1501 AAAGGGCTCGTGGAGAGACAACTCAACAGCAGGAGGATGCTTCTGGGC 1560
 QY 1561 AAATGGATCCCTATGGCCCTGCTCGGCACATGTGGTGGGGCGTGCAGCTGGCCAGG 1620
 Db 1561 AAATGGATCCCTATGGCCCTGCTCGGCACATGTGGTGGGGCGTGCAGCTGGCCAGG 1620
 QY 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGGCAGTACTCGAGGGAGTGGAGGTG 1680
 Db 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGGCAGTACTCGAGGGAGTGGAGGTG 1680
 QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCTCCCGAAAGAGTTCGGG 1740
 Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCTCCCGAAAGAGTTCGGG 1740
 QY 1741 GAGGAGCAGTGTGAGGCTTTCACCGCTTACACAGCAGCACCACCGGCTCAGCTCTGCC 1800
 Db 1741 GAGGAGCAGTGTGAGGCTTTCACCGCTTACACAGCAGCACCACCGGCTCAGCTCTGCC 1800
 QY 1801 GTGGATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTCAAGCTCATCTGC 1860
 Db 1801 GTGGATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTCAAGCTCATCTGC 1860
 QY 1861 CGAGCAATGGCAGTGGCTTCTATGTGTGGCACCACAGGTGGTGGAGCGGACGCTG 1920
 Db 1861 CGAGCAATGGCAGTGGCTTCTATGTGTGGCACCACAGGTGGTGGAGCGGACGCTG 1920
 QY 1921 TGCTCTCTCTGACTCCACCTCCCTGTGTGTCCAAAGCAAGTCAAGGCTGGCTGTGAT 1980
 Db 1921 TGCTCTCTCTGACTCCACCTCCCTGTGTGTCCAAAGCAAGTCAAGGCTGGCTGTGAT 1980
 QY 1981 GGGAACTGGGTCCAGAGAGATTCAGCAAGTGTGGGTGTGGGGGAGACAATAAG 2040
 Db 1981 GGGAACTGGGTCCAGAGAGATTCAGCAAGTGTGGGTGTGGGGGAGACAATAAG 2040
 QY 2041 AGCTGCAAGAAGTGAAGTGTGACTTCCACCAAGCCATGCATGGCTACAAATTCGTGGTG 2100
 Db 2041 AGCTGCAAGAAGTGAAGTGTGACTTCCACCAAGCCATGCATGGCTACAAATTCGTGGTG 2100
 QY 2101 GCATCCCGAGGGCTTCAAGCATCAGCATCCGCGCAGCGGTTTCAAGGCGCTGATC 2160
 Db 2101 GCATCCCGAGGGCTTCAAGCATCAGCATCCGCGCAGCGGTTTCAAGGCGCTGATC 2160
 QY 2161 GGGGATGACAATCTACCTGGCTCTGAGAACAGCAAGCAAGTACCTGCTCAACGGGAT 2220
 Db 2161 GGGGATGACAATCTACCTGGCTCTGAGAACAGCAAGCAAGTACCTGCTCAACGGGAT 2220
 QY 2221 TTGCTGGTGTGCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTCGGTACAGC 2280
 Db 2221 TTGCTGGTGTGCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTCGGTACAGC 2280
 QY 2281 GGACGGGACAGCGGTGGAGAGCTGAGGCTTCCCGGCCATCTGGAGCGGCTGACC 2340
 Db 2281 GGACGGGACAGCGGTGGAGAGCTGAGGCTTCCCGGCCATCTGGAGCGGCTGACC 2340
 QY 2341 GTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
 Db 2341 GTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
 QY 2401 CCCAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTGTC 2460
 Db 2401 CCCAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGGACCCCGGGGACCTCTGTGTC 2460
 QY 2461 TTGCACAAGAGCTCTCAGCTCTCCCAAGAGTGGAGCGGAGCAGGAGCCGCT 2520
 Db 2461 TTGCACAAGAGCTCTCAGCTCTCCCAAGAGTGGAGCGGAGCAGGAGCCGCT 2520
 QY 2521 GCACGTGGGTGGCTGGCAGCTGGGGCGGTGCTCCCGGAGGTGGCGAGTGGGCTGCAG 2580
 Db 2521 GCACGTGGGTGGCTGGCAGCTGGGGCGGTGCTCCCGGAGGTGGCGAGTGGGCTGCAG 2580

QY 2581 AAGCGCGGTGGACTGTCCGGGTCCGGCGGAGCGACGGTCCCTGCTGTGA 2640
 Db 2581 AAGCGCGGTGGACTGTCCGGGTCCGGCGGAGCGACGGTCCCTGCTGTGA 2640
 QY 2641 GCCCATCGCCGCTGGAGACACAAAGCTTCGCGGAGCGCTTCCCGACCTGGGAGCT 2700
 Db 2641 GCCCATCGCCGCTGGAGACACAAAGCTTCGCGGAGCGCTTCCCGACCTGGGAGCT 2700
 QY 2701 GCCTGTCACTCTGTCTCAAGAGCTTCGCGCGGGATTTTCAGAGCGCTCAGTCAAGTGT 2760
 Db 2701 GCCTGTCACTCTGTCTCAAGAGCTTCGCGCGGGATTTTCAGAGCGCTCAGTCAAGTGT 2760
 QY 2761 GTGGCCACGAGGCGGCTGTGTGGCGGAGCGAGTGTCAAGTGTGACCGAGCGGAG 2820
 Db 2761 GTGGCCACGAGGCGGCTGTGTGGCGGAGCGAGTGTCAAGTGTGACCGAGCGGAG 2820
 QY 2821 GAGCTGGACTTCTGCTGCTGAGGCGGTGCTGA 2853
 Db 2821 GAGCTGGACTTCTGCTGCTGAGGCGGTGCTGA 2853

RESULT 2
 AX342635
 LOCUS AX342635
 DEFINITION Sequence 32 from Patent WO0198468.
 ACCESSION AX342635
 VERSION AX342635.1 GI:18152032
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Yue.H., Elliott.V.S., Gandhi.A.R., Lal.P., Au-Young.J.,
 Tribouley.C.M., Deleage.A.M., Baughn.M.R., Nguyen.D.P., Lee.E.A.,
 Hafalla.A., Khan.F.A., Walla.N.K., Yao.M.G., Lu.D.A., Patterson.C.,
 Tang.Y.T., Walsh.R.T., Azimzal.Y., Ramkumar.J., Xu.Y. and Reddy.R.
 Patent: WO 0198468-A 32 27-DEC-2001;
 Incyte Genomics, Inc. (US)

FEATURES
 Location/Qualifiers
 source 1..2930
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7473089CB1"

BASE COUNT 529 a 958 c 931 g 512 t

ORIGIN

Query Match 98.8%; Score 2817.6; DB 6; Length 2930;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2846; Conservative 0; Mismatches 4; Indels 6; Gaps 2;

QY 1 ATGCTTCTCTGGGATCTTAACCTGGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
 Db 75 ATGCTTCTCTGGGATCTTAACCTGGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 134
 QY 61 CCAGAGCGGAGGTAGTCTTCCCATTCGACTTGGACCGGACATTAAGCGCGCGCTAC 120
 Db 135 CCAGAGCGGAGGTAGTCTTCCCATTCGACTTGGACCGGACATTAAGCGCGCGCTAC 194
 QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCAGAGCATTT 180
 Db 195 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCAGAGCATTT 254
 QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTAGTTCCTGGCTCCCGCTTCC 240
 Db 255 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTAGTTCCTGGCTCCCGCTTCC 314
 QY 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGCTCAGCGGGGCTCTTACAGCTGGAGCG 300
 Db 315 ACTGAGCATCTGGGCTCCCGCTCCAGGGCTCAGCGGGGCTCTTACAGCTGGAGCG 374
 QY 301 TGCTTCTATTCTGGGAGCTGAACCGGAGCTGCTTCTGCTGCTGTGAGGCTGTGC 360

|||||
Db 375 TGCTTCTATTCTGGGACGTGAACGGCGGACCGGACTCGTTGCTGCTGTGAGCGCTGTGC 434
QY 361 GGGGGCTCCGCGAGCGCTTTGGCTACCGAGGCGCGAGTAGTGTCAATTAGCGCGCTGCGCC 420
Db 435 GGGGGCTCCGCGAGCGCTTTGGCTACCGAGGCGCGAGTAGTGTCAATTAGCGCGCTGCGCC 494
QY 421 AATGCTAGCGCGCGGGCGGCGAGCGACAGACGAGGCGCCACACCTTCTCAGCGCGCG 480
Db 495 AATGCTAGCGCGCGGGCGGCGAGCGAACAGCGAGGCGGCACACCTTCTCAGCGCGCG 554
QY 481 GGTGTTCCGGCGGGCGCTTCCGAGAGACCCACCTCTCGTTCGGGGTGGTCTCGGGCTG 540
Db 555 GGTGTTCCGGCGGGCGCTTCCGAGAGACCCCACTCTCGTTCGGGGTGGTCTCGGGCTG 614
QY 541 AACCCCGCCATCTTACGGGCGCTTGAGACCTTTACAAGCGCGGGCGGGCGGCTTCGGGGAG 600
Db 615 AACCCCGCCATCTTACGGGCGCTTGAGACCTTTACAAGCGCGGGCGGGCGGCTTCGGGGAG 674
QY 601 AGTGTAGCGCGCGAGCTGTGGCGCGCGCAAGCTTTGCTGTCTATCCCGCGGTACGTG 660
Db 675 AGTGTAGCGCGCGAGCTGTGGCGCGCGCAAGCTTTGCTGTCTATCCCGCGGTACGTG 734
QY 661 GAGACGCTGTGTGCGGCGAGTCAATGCTCAAGTTTCCACGCGCGGGCGGACCTTGGAAAT 720
Db 735 GAGACGCTGTGTGCGGCGAGTCAATGCTCAAGTTTCCACGCGCGGGCGGACCTTGGAAAT 794
QY 721 TATCTGTAGCGCTGTGGCAAGCGCGCGGAGCTCTACCGCCATCCACATCTCTCAAC 780
Db 795 TATCTGTAGCGCTGTGGCAAGCGCGCGGAGCTCTACCGCCATCCACATCTCTCAAC 854
QY 781 CCCATCAACATCTGTGTGTTGTTCAAGTGTCTTCTTAGAGATCTGACTCCGGGCGGCAAG 840
Db 855 CCCATCAACATCTGTGTGTTGTTCAAGTGTCTTCTTAGAGATCTGACTCCGGGCGGCAAG 914
QY 841 GTCACCGGCAATGGGCGCTGACCGTGGGCAACTTCTGTGGCTGGGAGAGAAAGCTCAAC 900
Db 915 GTCACCGGCAATGGGCGCTGACCGTGGGCAACTTCTGTGGCTGGGAGAGAAAGCTCAAC 974
QY 901 AAAGTGAGTGACAGACCCCGAGTACTGGGACACTGCCATCTCTTACACGAGCAGGAC 960
Db 975 AAAGTGAGTGACAGACCCCGAGTACTGGGACACTGCCATCTCTTACACGAGCAGGAC 1034
QY 961 CTGTGTGAGGCGACACCTGTGACACCGTGGGCAATGGTGTGAGTGGGTACCATGTGTGAC 1020
Db 1035 CTGTGTGAGGCGACACCTGTGACACCGTGGGCAATGGTGTGAGTGGGTACCATGTGTGAC 1094
QY 1021 CCCAAGAGAAAGTCTCTCTTATTGAGGACGATGGGCTTCCATCAGCTTCCACCACTGCC 1080
Db 1095 CCCAAGAGAAAGTCTCTCTTATTGAGGACGATGGGCTTCCATCAGCTTCCACCACTGCC 1154
QY 1081 CACGAGCTGGGCGAGCTGTTCAACATGCGCCATGACAAATGTGAAGTCTGTGAGGAGTG 1140
Db 1155 CACGAGCTGGGCGAGCTGTTCAACATGCGCCATGACAAATGTGAAGTCTGTGAGGAGTG 1214
QY 1141 TTTGGGAAGCTCCGAGCGCAACACATGATGTCCCGAGCCCTCATCCAGATCGACCGTGCC 1200
Db 1215 TTTGGGAAGCTCCGAGCGCAACACATGATGTCCCGAGCCCTCATCCAGATCGACCGTGCC 1274
QY 1201 AACCCCTGTGACGCTGCAGTGTCTGCCATCATACCGACTTCTCTGGACAGCGGGCAGCGT 1260
Db 1275 AACCCCTGTGACGCTGCAGTGTCTGCCATCATACCGACTTCTCTGGACAGCGGGCAGCGT 1334
QY 1261 GACTGCTCTCTGGACCAACCCAGAACGCCATCTCCCTGCGCGAGGATCTCGCGGGCGCC 1320
Db 1335 GACTGCTCTCTGGACCAACCCAGAACGCCATCTCCCTGCGCGAGGATCTCGCGGGCGCC 1394
QY 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCTGGCTTCCAAAGCCCTGTGCT 1380
Db 1395 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCTGGCTTCCAAAGCCCTGTGCT 1454
QY 1381 TACATGCACTACTGCAACCAAGCTGTGGTGCACCGGGAAGGCGAAGGAGCAGATGTTGTC 1440
|||||

Db 1455 TACATGCACTACTGCAACCAAGCTGTGGTGTGCACCGGGAAGGCCAACGGAGACAGATGTTGTC 1514
QY 1441 CAGACCCGCCACTTTCCTCTGGCGCGATGGCACCAAGCTGTGGCGAGGGCAAGCTTGGCTC 1500
Db 1515 CAGACCCGCCACTTTCCTCTGGCGCGATGGCACCAAGCTGTGGCGAGGGCAAGCTTGGCTC 1574
QY 1501 AAAGGGGCTCGGTGGAGAGACAACTCAACAGACAGGGTGGATGGTTCCTGGGCC 1560
Db 1575 AAAGGGGCTCGGTGGAGAGACAACTCAACAGACAGGGTGGATGGTTCCTGGGCC 1634
QY 1561 AAATGGGATCCCTATGGCCCTCTCGCGCACATGTGTGGGGCGTGCAGCTGGCCAGG 1620
Db 1635 AAATGGGATCCCTATGGCCCTCTCGCGCACATGTGTGGGGCGTGCAGCTGGCCAGG 1694
QY 1621 AGGCAAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1680
Db 1695 AGGCAAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1754
QY 1681 AAATACCGATCCCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGGAAGAGCTTCGG 1740
Db 1755 AAATACCGATCCCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGGAAGAGCTTCGG 1814
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACACCAAGCAACCGGCTCACTTCGCC 1800
Db 1815 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACACCAAGCAACCGGCTCACTTCGCC 1874
QY 1801 GTGGCATGGGTGCCCAAGTACTCCGGCGTGTCTCCCGGGACAGTGCAAGCTCAATGTC 1860
Db 1875 GTGGCATGGGTGCCCAAGTACTCCGGCGTGTCTCCCGGGACAGTGCAAGCTCAATGTC 1934
QY 1861 CGAGCAATGGGCACTGGCTTCTATGTGTGGCACCACAGGTGGTGGACCGGACG 1917
Db 1935 CGAGCAATGGGCACTGGCTTCTATGTGTGGCACCACAGGTGGTGGACCGGACG 1994
QY 1918 CTGTGCTCTCTGACTTCCACCTCCCTGCTGTGTCTCAAGCAAGTGCATCAAGCTGGCTGT 1977
Db 1995 CTGTGCTCTCTGACTTCCACCTCCCTGCTGTGTCTCAAGCAAGTGCATCAAGCTGGCTGT 2054
QY 1978 GATGGAACTGGGTGCCAAGAGAGATTCGACAAAGTGTGGGTGTGGGGAGACAGAT 2037
Db 2055 GATGGAACTGGGTGCCAAGAGAGATTCGACAAAGTGTGGGTGTGGGGAGACAGAT 2114
QY 2038 AAGAGCTGAAGAAGTGTGACTTCTCAGCAAGCCCATCATGCTGGCTTACAAATTCGCTG 2097
Db 2115 AAGAGCTGAAGAAGTGTGACTTCTCAGCAAGCCCATCATGCTGGCTTACAAATTCGCTG 2174
QY 2098 GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCGACGCGGTTTACAAAGGCTG 2157
Db 2175 GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCGACGCGGTTTACAAAGGCTG 2234
QY 2158 ATCGGGATGACAACTTACCTGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGG 2217
Db 2235 ATCGGGATGACAACTTACCTGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGG 2294
QY 2218 CATTTGCTGTGTGCGGCTGGAGCGGACCTGTGGTGAAGGGCAGTGTCTGTGGGGTAC 2277
Db 2295 CATTTGCTGTGTGCGGCTGGAGCGGACCTGTGGTGAAGGGCAGTGTCTGTGGGGTAC 2354
QY 2278 AGCGGACGCGGACAGGGTGGAGAGCTTCAGGCTTCCCGCCCATCTCTGGAGCGGCTG 2337
Db 2355 AGCGGACGCGGACAGGGTGGAGAGCTTCAGGCTTCCCGCCCATCTCTGTGGCGGCTG 2414
QY 2338 ACCTGGAGGTCTCTCTCGGTGGGGAAGATGACACCGCCCGCGGTCTGCTTCTTAT 2397
Db 2415 ACCTGGAGGTCTCTCTCGGTGGGGAAGATGACACCGCCCGCGGTCTGCTTCTTAT 2474
QY 2398 CTGCCCCAAGAGCTCGGGAGGACAGTCTCTCATCCCAAGAGACCCCGG ---GGACCC 2454
Db 2475 CTGCCCCAAGAGCTCGGGAGGACAGTCTCTCATCCCGCCAGCCCGCGGGAGAGCC 2534
QY 2455 TCTGTCTTGCACACAGCGTCTCAGCTCTCCAAACAGTGGAGCAGCGGACGACAGG 2514
Db 2535 TCTGTCTTGCACACAGCGTCTCAGCTCTCCAAACAGTGGAGCAGCGGACGACAGG 2594
|||||

```
QY 2515 CCCCTGACAGCTGGGTGGCTGGAGCTGGGGGCGGTGCTCCGGAGAGCTCGGAGTGGC 2574
Db 2595 CCCCTGACAGCTGGGTGGCTGGAGCTGGGGGCGGTGCTCCGGAGAGCTCGGAGTGGC 2654
QY 2575 CTGCGAGAGCGGGCGGTGGAGCTGGGGGCTTCGCGGGGAGCGCAGCGGTCCCTGCTGT 2634
Db 2655 CTGCGAGAGCGGGCGGTGGAGCTGGGGGCTTCGCGGGGAGCGCAGCGGTCCCTGCTGT 2714
QY 2635 GATGAGCGCCATCGCCCGCTGGAGACACAAGCTTCGGGGAGCGCTCCGCCACCTGGGAG 2694
Db 2715 GATGAGCGCCATCGCCCGCTGGAGACACAAGCTTCGGGGAGCGCTCCGCCACCTGGGAG 2774
QY 2695 CTCAGCGCTGTGTACACCTGTCTCAAGAGCTGGCGGCGGGGATTTACAGAGCGCTCACTC 2754
Db 2775 CTCAGCGCTGTGTACACCTGTCTCAAGAGCTGGCGGCGGGGATTTACAGAGCGCTCACTC 2834
QY 2755 AAGTGTGGGCGACGGAGCGGCTGTGCGCGGGAGCGCTTCGGGGAGCGCTCCGCCACCTGGGAG 2814
Db 2835 AAGTGTGGGCGACGGAGCGGCTGTGCGCGGGAGCGCTTCGGGGAGCGCTCCGCCACCTGGGAG 2894
QY 2815 CCCAGGAGCTGGACTTTCGGTCTCTGAGCGCGTGC 2850
Db 2895 CCCAGGAGCTGGACTTTCGGTCTCTGAGCGCGTGC 2930

RESULT 3
AX319860
LOCUS AX319860 2937 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 24 from Patent WO0183782.
ACCESSION AX319860
VERSION AX319860.1 GI:17901450
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Sudarsanam,S., Manning,S., Caenepeel,S. and
Payne,V.
TITLE Novel proteases
JOURNAL Patent: WO 0183782-A 24 08-NOV-2001;
Sugen, Inc. (US)
FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 531 a 958 c 936 g 512 t ;
ORIGIN
Query Match 98.8%; Score 2817.6; DB 6; Length 2937;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2849; Conservative 0; Mismatches 4; Indels 9; Gaps 2;
QY 1 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCGCGGGGCAACCGCTGGAGGCTCTGAG 60
Db 76 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCGCGGGGCAACCGCTGGAGGCTCTGAG 135
QY 61 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGACCGCGACATTAAACGGCGCGCGGTAC 120
Db 136 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGACCGCGACATTAAACGGCGCGCGGTAC 195
QY 121 TACTGGCGGGGTCGCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGAGCAATT 180
Db 196 TACTGGCGGGGTCGCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGAGCAATT 255
QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCCTGGGTCGCGGCTTCTCC 240
Db 256 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCCTGGGTCGCGGCTTCTCC 315
QY 241 ACTGAGCATCTGGGGTCCCTCCAGGGGCTCACCAGGGGCTCTCAGACCTCGGAGGC 300
Db 316 ACTGAGCATCTGGGGTCCCTCCAGGGGCTCACCAGGGGCTCTCAGACCTCGGAGGC 375
```

```
QY 301 TCGTTCTATTCTGGGAGCTGAAGCGCGAGCGGACTCGTTCCGCTGCTGTGAGCTGTGC 360
Db 376 TCGTTCTATTCTGGGAGCTGAAGCGCGAGCGGACTCGTTCCGCTGCTGTGAGCTGTGC 435
QY 361 GGGGGGCTCCGGGAGCGCTTTGGCTACCGAGGCGCGAGTATGATCATAGCCGCTGCC 420
Db 436 GGGGGGCTCCGGGAGCGCTTTGGCTACCGAGGCGCGAGTATGATCATAGCCGCTGCC 495
QY 421 AATGCTAGCGCGCGCGCGCGAGCGCAACAGCGAGGCGCACACCTTCTCCAGCGCGG 480
Db 496 AATGCTAGCGCGCGCGCGCGAGCGCAACAGCGAGGCGCACACCTTCTCCAGCGCGG 555
QY 481 GGTGTTCGGGGGGGGGCTTCGCGAGACCCACCTCTCGCTCGGGGTGGGCTTCGGGTG 540
Db 556 GGTGTTCGGGGGGGGGCTTCGCGAGACCCACCTCTCGCTCGGGGTGGGCTTCGGGTG 615
QY 541 AACCCCGGCATCTACGGGCGCTTGAGCCCTTACAGCGCGGCGGGGCTTCGGGAG 600
Db 616 AACCCCGGCATCTACGGGCGCTTGAGCCCTTACAGCGCGGCGGGGCTTCGGGAG 675
QY 601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCCAAAGCGTTTCGTGTCTATCCCGCGGTAG 660
Db 676 AGTCGTAGCGCGCGAGGTCTGGGCGCGCCAAAGCGTTTCGTGTCTATCCCGCGGTAG 735
QY 661 GAGACGCTGGTGGTGGCGGAGAGTCAATGCTCAAGTTCACGCGGCGGAGCTTGAACAT 720
Db 736 GAGACGCTGGTGGTGGCGGAGAGTCAATGCTCAAGTTCACGCGGCGGAGCTTGAACAT 795
QY 721 TATCTGCTGACGCTCTGCGCAACGCGCGGAGTCTTACCGCCCATCCAGCATCTCAAC 780
Db 796 TATCTGCTGACGCTCTGCGCAACGCGCGGAGTCTTACCGCCCATCCAGCATCTCAAC 855
QY 781 CCCATCAACATCGTTGGTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCAAG 840
Db 856 CCCATCAACATCGTTGGTCAAGTGTCTTCTTAGAGATCGTACTCCGGGCAAG 915
QY 841 GTCACCGGCAATGCGGCGCTGACGCTGGCAACTTCTGCTGGCAGAGAAGCTGAAC 900
Db 916 GTCACCGGCAATGCGGCGCTGACGCTGGCAACTTCTGCTGGCAGAGAAGCTGAAC 975
QY 901 AAAGTAGTAGCAAGACCGCGAGTCTGGGAGACTGCGATCTTCCACGAGCAAG 960
Db 976 AAAGTAGTAGCAAGACCGCGAGTCTGGGAGACTGCGATCTTCCACGAGCAAG 1035
QY 961 CTGTGTGGAGCCACCACTGTGACACCGCTGGGATGGCTGATGTGGGTACCATGTG 1020
Db 1036 CTGTGTGGAGCCACCACTGTGACACCGCTGGGATGGCTGATGTGGGTACCATGTG 1095
QY 1021 CCCAAGAGAGCTGCTGTGTCATTGAGGAGCATGGGCTTCCATCAGCCTTCCACCA 1080
Db 1096 CCCAAGAGAGCTGCTGTGTCATTGAGGAGCATGGGCTTCCATCAGCCTTCCACCA 1155
QY 1081 CAGGAGCTGGGCGCAGTGTTCACATGCCCATGCAATGTGAAGTCTGTGAGAGGTG 1140
Db 1156 CAGGAGCTGGGCGCAGTGTTCACATGCCCATGCAATGTGAAGTCTGTGAGAGGTG 1215
QY 1141 TTTGGAGAGCTCCGAGCAACCAACATGATGTCCCGACCCCTCATCCAGATCGACCTG 1200
Db 1216 TTTGGAGAGCTCCGAGCAACCAACATGATGTCCCGACCCCTCATCCAGATCGACCTG 1275
QY 1201 AACCCCTGTGCTGAGCTGCGAGTGTGCCATCATACCGAGTTCCTGGACAGGGCA 1260
Db 1276 AACCCCTGTGCTGAGCTGCGAGTGTGCCATCATACCGAGTTCCTGGACAGGGCA 1335
QY 1261 GACTGCTCTCTGGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGTGCGGGG 1320
Db 1336 GACTGCTCTCTGGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGTGCGGGG 1395
QY 1321 AGCTACACCTCAGCCAGGAGTGGGCTTTTGGGTGGGCTTCCAGCCCTGTCT 1380
Db 1396 AGCTACACCTCAGCCAGGAGTGGGCTTTTGGGTGGGCTTCCAGCCCTGTCT 1455
```

```
QY 1381 TACATCAGTACTGACCAAGCTGTGTCACCGGAAAGCCCAAGGACACAGTGTGTC 1440
|||||
Db 1456 TACATCAGTACTGACCAAGCTGTGTCACCGGAAAGCCCAAGGACAGTGTGTC 1515
|||||
QY 1441 CAGACCCGCACTTCCCTGGGCCGATGACACAGCTGTGGCGAGGGCAAGCTCTGCCTC 1500
|||||
Db 1516 CAGACCCGCACTTCCCTGGGCCGATGACACAGCTGTGGCGAGGGCAAGCTCTGCCTC 1575
|||||
QY 1501 AAAGGGCCCTGCGTGAGAGACACACCTCAACACACACAGGTCGATGCTTCTGGGCC 1560
|||||
Db 1576 AAAGGGCCCTGCGTGAGAGACACACCTCAACACACACAGGTCGATGCTTCTGGGCC 1635
|||||
QY 1561 AAATGGATCCCTATGGCCCTCTCGCCACATGTGTGGGGCGCTGACGTGGCCAGG 1620
|||||
Db 1636 AAATGGATCCCTATGGCCCTCTCGCCACATGTGTGGGGCGCTGACGTGGCCAGG 1695
|||||
QY 1621 AGGAGTGCACCAACCCACCCCTGCCAAGCGGGGCAAGTACTGCGAGGAGTGGGGTG 1680
|||||
Db 1696 AGGAGTGCACCAACCCACCCCTGCCAAGCGGGGCAAGTACTGCGAGGAGTGGGGTG 1755
|||||
QY 1681 AAATACCGATCCCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCTCCGGAAGAGCTTCCGG 1740
|||||
Db 1756 AAATACCGATCCCTGCAATCTGGAGCCCTGCCCCAGCTCAGCCTCCGGAAGAGCTTCCGG 1815
|||||
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTACACCAAGACCAACCGGCTCAGCTTCGCC 1800
|||||
Db 1816 GAGGAGCAGTGTGAGGCTTTCAACGGCTACACCAAGACCAACCGGCTCAGCTTCGCC 1875
|||||
QY 1801 GTGCATGGGTGCCAAGTACTCGGGGTGTCTCCCGGACAAAGTGCAAGCTCATCTGC 1860
|||||
Db 1876 GTGCATGGGTGCCAAGTACTCGGGGTGTCTCCCGGACAAAGTGCAAGCTCATCTGC 1935
|||||
QY 1861 CGAGCCAATGGCACTGCTACTTCTATGTCTGCGCACCCAA---GGTGTGGACGGCACG 1917
|||||
Db 1936 CGAGCCAATGGCACTGCTACTTCTATGTCTGCGCACCCAAAGGTGTGGACGGCACG 1995
|||||
QY 1918 CTGTGCTCTCTGACTCCACTCGCTGTGTCTCAAGGCAAGTGCATCAAGGCTGCTGT 1977
|||||
Db 1996 CTGTGCTCTCTGACTCCACTCGCTGTGTCTCAAGGCAAGTGCATCAAGGCTGCTGT 2055
|||||
QY 1978 CATGGAACTGGGCTCCAAAGAGAGATCGACAAGTGGGTGTGTGGGGGAGACAAT 2037
|||||
Db 2056 CATGGAACTGGGCTCCAAAGAGAGATTCGACAGTGGGTGTGTGGGGGAGACAAT 2115
|||||
QY 2038 AAGAGCTGCAAGAAGTGTGACTGCTCTTCCAAAGCCCATGATGCTACAAATTCGTG 2097
|||||
Db 2116 AAGAGCTGCAAGAAGTGTGACTGCTCTTCCAAAGCCCATGATGCTACAAATTCGTG 2175
|||||
QY 2098 GTGGCCATCCCGCAGGGGCTCAAGCATCGACATCCCGCAGCGGTTACAAGGGCTG 2157
|||||
Db 2176 GTGGCCATCCCGCAGGGGCTCAAGCATCGACATCCCGCAGCGGTTACAAGGGGCTG 2235
|||||
QY 2158 ATCGGGATGACAATCTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAAGCGG 2217
|||||
Db 2236 ATCGGGATGACAATCTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAAGCGG 2295
|||||
QY 2218 CATTTCTGTGTGTGCGGGTGGAGCGGACCTGTGTGTGAAGGCACTCTGTCGGGTAC 2277
|||||
Db 2296 CATTTCTGTGTGTGCGGGTGGAGCGGACCTGTGTGTGAAGGCACTCTGTCGGGTAC 2355
|||||
QY 2278 AGCGCACGGGCACAGCGGTGGAGGCTGACAGCTTCCGGCCCATCTCTGGAGCGGCTG 2337
|||||
Db 2356 AGCGCACGGGCACAGCGGTGGAGGCTGACAGCTTCCGGCCCATCTCTGGAGCGGCTG 2415
|||||
QY 2338 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2397
|||||
Db 2416 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2475
|||||
QY 2398 CTGCCCAAAGAGCCTCGGGAGGACAAGTCTCTCATGCCAAGAGACCCCGG-----GGA 2451
|||||
Db 2476 CTGCCCAAAGAGCCTCGGGAGGACAAGTCTCTCATGCCCGACCCCGGGGAGAGGA 2535
|||||
QY 2452 CCCTCTGTCTTTCACACACAGCGTCTCTCAGCCTCTCCACACAGGTGGAGCAGCGGACGAC 2511
|||||
```

```
Db 2536 CCCTCTGTCTTGCACAACAGCGCTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGAC 2595
|||||
QY 2512 AGGCCCCCTGCACGCTGGGTGGCTGGCAGCTGGGGGCCGTGCTCCGGAGCTCGGCGAGT 2571
|||||
Db 2596 AGGCCCCCTGCACGCTGGGTGGCTGGCAGCTGGGGGCCGTGCTCCGGAGCTCGGCGAGT 2655
|||||
QY 2572 GGCTGCAAGAAAGCGGGGCTGACTGTGGGGGTCTCCCGGGGAGCGACAGTTCCTTCC 2631
|||||
Db 2656 GGCTGCAAGAAAGCGGGGCTGACTGTCCGGGGTCTCCCGGGGAGCGACAGCTTCCCTGCC 2715
|||||
QY 2632 TGTGATGAGCCCATCGGCGCGCTGGAGACACAGCCTTGGGGGAGCCCTGCCACCTGG 2691
|||||
Db 2716 TGTGATGAGCCCATCGGCGCGCTGGAGACACAGCCTTGGGGGAGCCCTGCCACCTGG 2775
|||||
QY 2692 GAGCTCAGCGCTGTGCTCAGCCTGTCTCAAGAGCTGCGCGCGGGGATTCAGAGCGGCTCA 2751
|||||
Db 2776 GAGCTCAGCGCTGTGCTCAGCCTGTCTCAAGAGCTGCGCGCGGGGATTCAGAGCGGCTCA 2835
|||||
QY 2752 CTCAGTGTGTGGGCGACGAGCGCGCTGCTGCCCGGGACCAAGTGCACATTCGACCGC 2811
|||||
Db 2836 CTCAGTGTGTGGGCGACGAGCGCGCTGCTGCCCGGGACCAAGTGCACATTCGACCGC 2895
|||||
QY 2812 AAGCCCCAGGAGCTGACTTCTGCTGAGGCGCTGCTGA 2853
|||||
Db 2896 AAGCCCCAGGAGCTGACTTCTGCTGAGGCGCTGCTGA 2937
|||||

RESULT 4
AP002986 157963 bp DNA linear p: 08-MAR-2002
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-121M22,
DEFINITION complete sequence.
ACCESSION AP002986
VERSION AP002986.2 GI:19263031
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-121M22.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Eumleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H.
1
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seon,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seon,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Mar 7, 2002 this sequence version replaced gi:11559301.
FEATURES
source
1..157963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-121M22"

BASE COUNT 39742 a 36722 c 37247 g 44252 t
ORIGIN

Query Match 33.7%; Score 960.4; DB 9; Length 157963;
Best Local Similarity 98.4%; Pred. No. 6.7e-163;
Matches 970; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGCATCTAACCCTGCTTTCGCGGGCGGCAACCGCTGGAGGCTCTGAG 60
|||||
Db 134115 ATGCTTCTGCTGGGCATCTAACCCTGCTTTCGCGGGCGGCAACCGCTGGAGGCTCTGAG 134174
|||||
```


Assembly program: Phrap: version 0. 960731
 Consensus quality: 161899 bases at least Q40.
 Consensus quality: 167402 bases at least Q30.
 Consensus quality: 165801 bases at least Q20
 Insert size: 177000; agarose-fp
 Insert size: 171405; sum-of-contigs
 Quality coverage: 4.1 in Q20 bases; agarose-fp
 Quality coverage: 4.2 in Q20 bases; sum-of-con-

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

*	1	21: contig of 21 bp in length
*	12	121: gap of 100 bp
*	122	1343: contig of 1222 bp in length
*	1344	1443: gap of 100 bp
*	1444	2636: contig of 1193 bp in length
*	2637	2736: gap of 100 bp
*	2737	4052: contig of 1316 bp in length
*	4053	4152: gap of 100 bp
*	4153	8044: contig of 3892 bp in length
*	8045	8144: gap of 100 bp
*	8145	14949: contig of 6805 bp in length
*	14950	15049: gap of 100 bp
*	15050	23535: contig of 8496 bp in length
*	23536	23635: gap of 100 bp
*	23636	30261: contig of 6626 bp in length
*	30262	30361: gap of 100 bp
*	30362	40941: contig of 10580 bp in length
*	40942	41041: gap of 100 bp
*	41042	54893: contig of 13852 bp in length
*	54894	54993: gap of 100 bp
*	54994	68888: contig of 13895 bp in length
*	68889	69988: gap of 100 bp
*	69989	83702: contig of 14714 bp in length
*	83703	83802: gap of 100 bp
*	83803	98393: contig of 14591 bp in length
*	98394	98493: gap of 100 bp
*	98494	117356: contig of 18863 bp in length
*	117357	117456: gap of 100 bp
*	117457	140589: contig of 23133 bp in length
*	140590	140689: gap of 100 bp
*	140690	172905: contig of 32216 bp in length.

FEATURES
SOURCE

```
1. .1/2903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-211H6"
/clone_lib="RPC1-11 Human Male BAC"
1. .21
```

misc_feature

```

/note="assembly_fragment
clone_end:T7
vector_side:right"

```

misc_feature

```
1221.1349
/note="assembly_fragment"
1444.2636
/note="assembly_v fragment"
```

misc_feature

```
2737: 4052
/note="assembly_fragment"
4153: .8044
/note="assembly_v fragment"
```

misc feature

```
0143. .14343
/note="assembly_fragment"
15050. .23535
/note="assembly_fragment"
```

misc_feature

TOZOC: .06067

		/note="assembly_fragment clone_end:SP6 vector_side:right" 30362..40941 /note="assembly_fragment" 41042..54893 /note="assembly_fragment" 54994..68888 /note="assembly_fragment" 68989..83702 /note="assembly_fragment" 83803..98393 /note="assembly_fragment" 98494..117356 /note="assembly_fragment" 117457..140589 /note="assembly_fragment" 140690..172905 /note="assembly_fragment"
BASE COUNT	44360 a 41333 c 39516 g 46194 t 1502 others	
ORIGIN		
	Query Match 33.7%; Score 960.4; DB 2; Length 172905; Best Local Similarity 98.4%; Pred. No. 6.6e-163; Matches 970; Conservative 0; Mismatches 16; Indels 0; Gaps 0;	
Qy	1 ATGCTTCGTGGGCATCCTAACCTGCTTTCCCGCGGCGCAACCGCTGGAGGCTCTTGAG 60	
Dd	79583 ATGCTTCGTGGGCATCCTAACCTGCTTTCCCGCGGCGCAACCGCTGGAGGCTCTTGAG 79524	
Qy	61 CCAGAGCGGGAGTGTGCTTTCCCATCCGACTGGACCGCGGACATTAAACGGCGCGGCTAC 120	
Dd	79523 CCAGAGCGGGAGTGTGCTTTCCCATCCGACTGGACCGCGGACATTAAACGGCGCGGCTAC 79464	
Qy	121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTTTCAGATCACAGCATTT 180	
Dd	79463 TACTGGCGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTTTCAGATCACAGCATTT 79404	
Qy	181 CAGGAGGACTTTTACTTACACTGACGCCGATGCTCAGTTCTTGGCTCCCCGCTTTGCC 240	
Dd	79403 CAGGAGGACTTTTACTTACACTGACGCCGATGCTCAGTTCTTGGCTCCCCGCTTTGCC 79344	
Qy	241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTACCGGGGGCTCTTCAGACCITGGACGCG 300	
Dd	79343 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTACCGGGGGCTCTTCAGACCITGGACGCG 79284	
Qy	301 TCCTTCTATTCTGGGAGTGAACCCGAGCGCGACTCGTTGCGTGTGTGAGCCTGTGCG 360	
Dd	79283 TGCTTCTATTCTGGGAGTGAACCCGAGCGCGACTCGTTGCGTGTGTGAGCCTGTGCG 79224	
Qy	361 GGGGGCTCCCGGGAGCCTTTTGCTTACGAGCGCGAGTAGTGTTCATTAGCCCGTGTGCC 420	
Dd	79223 GGGGGCTCCCGGGAGCCTTTTGCTTACGAGCGCGAGTAGTGTTCATTAGCCCGTGTGCC 79164	
Qy	421 AATGTAGCGCGCGGCGCGCAGCGCACACAGCCAGCGGCACACTTCTCCAGCGCGCGG 480	
Dd	79163 AATGTAGCGCGCGGCGCGCAGCGCACACAGCCAGCGGCACACTTCTCCAGCGCGCGG 79104	
Qy	481 GGTTGTCGGGGCGGGCCTTCCGGAGACCCCACTCTCGCTGCGGGGTGGCCTCGGGCTGG 540	
Dd	79103 GGTTGTCGGGGCGGGCCTTCCGGAGACCCCACTCTCGCTGCGGGGTGGCCTCGGGCTGG 79044	
Qy	541 AACCCCGGCATCTTACGGGCGCTGACCCCTTACAAGCGCGCGCGGGCGGGCTTCGGGGAG 600	
Dd	79043 AACCCCGGCATCTTACGGGCGCTGACCCCTTACAAGCGCGCGCGGGCGGGCTTCGGGGAG 78984	
Qy	601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCAAACGCTTTCGTCTCTATCCCGCGGTACGTTG 660	
Dd	78983 AGTCGTAGCGCGCGAGGTCTGGGCGCGCAAACGCTTTCGTCTCTATCCCGCGGTACGTTG 78924	
Qy	661 GAGACGCTGGTGGTCCGGACGAGTGCAATGGTCAAGTTCCACGGCGCGGACCTGGCAAT 720	
Dd	78923 GAGACGCTGGTGGTCCGGACGAGTGCAATGGTCAAGTTCCACGGCGCGGACCTGGCAAT 78864	


```

QY 721 TATCTGCTGACGCTGCTGGCAACGGCGGAGCTACTACCGCATCCAGCATCCTCAAC 780
Db 78863 TATCTGCTGACGCTGCTGGCAACGGCGGAGCTACTACCGCATCCAGCATCCTCAAC 78804
QY 781 CCATCAACATGTTGTGGTCAAGTGCTGCTTCTTAGAGATCGTGACTCGGCGCCCAAG 840
Db 78803 CCATCAACATGTTGTGGTCAAGTGCTGCTTCTTAGAGATCGTGACTCGGCGCCCAAG 78744
QY 841 GTCACGGCAATGGGCGCTGACGCTGGCAACTTCTGCTGGAGAGAAAGCTGAAC 900
Db 78743 GTCACGGCAATGGGCGCTGACGCTGGCAACTTCTGCTGGAGAGAAAGCTGAAC 78684
QY 901 AAGTGTAGTGAACACCCGAGTACTGGGACACTGCGATCTCTTACCAGCAGGAC 960
Db 78683 AAGTGTAGTGAACACCCGAGTACTGGGACACTGCGATCTCTTACCAGCAGGAC 78624
QY 961 CTGTGTGAGGACCAACCTGTGACAC 986
Db 78623 AGTTGATCTGCGTCACTTTGCACCC 78598

RESULT 6
AC023429 170682 bp DNA linear HTG 10-APR-2001
LOCUS Homo sapiens chromosome 11 clone RP11-121M22, WORKING DRAFT
DEFINITION SEQUENCE, 7 unordered pieces.
AC023429
AC023429.19 GI:13569974
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ABOLA, A.P., Bruno, D., Conn, L., Della Rosa, M., Faulkner, D.,
Fiederspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,
Mao, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M.,
Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D.,
Southwick, A.M., Webb, C., Wilhelmy, J., Yu, S. and Davis, R.W.
Unpublished
2 (bases 1 to 170682)
Bruno, D., Conn, L., Della Rosa, M., Faulkner, D., Federspiel, N.,
Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelmy, J.,
Yu, S. and Davis, R.W.
Direct Submission
Submitted (14-FEB-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
On Apr 10, 2001 this sequence version replaced gi:13562078.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development
Center
Center code: SDSTDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: 837
Center clone name: RP11-121M22
----- Summary Statistics
Sequencing Vector: M13mp18; X02513
Chemistry: Dye-primer; 12% of reads
Assembly: Dye-terminator Big Dye; 86% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165770 bases at least Q40
Consensus quality: 167255 bases at least Q30
Consensus quality: 167918 bases at least Q20
Insert size: 172423; agarose-fp
Insert size: 170082; sum-of-contents
Quality coverage: 9.0x in Q20 bases; agarose-fp
Quality coverage: 9.1x in Q20 bases; sum-of-contents.
* NOTE: This is a 'working draft' sequence. It currently

```

```

* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1856: contig of 1856 bp in length
* 1857 1956: gap of unknown length
* 1957 7065: contig of 5109 bp in length
* 7066 7165: gap of unknown length
* 7166 21089: contig of 13924 bp in length
* 21090 21189: gap of unknown length
* 21190 34145: contig of 12956 bp in length
* 34146 48562: gap of unknown length
* 48563 48663: contig of 14317 bp in length
* 48663 109179: gap of unknown length
* 109180 109279: contig of 60517 bp in length
* 109280 170682: gap of unknown length
* 170682 61403: contig of 61403 bp in length.

```

FEATURES

Source

```

1. 170682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-121M22"
/clone_lib="RPC1 human BAC library 11"
1. 1856
/note="assembly_name:Contig25"
1957..7065
/note="assembly_name:Contig26"
7166..21089
/note="assembly_name:Contig27"
21190..34145
/note="assembly_name:Contig28"
34246..48562
/note="assembly_name:Contig29"
clone_end:SP6"
48663..109179
/note="assembly_name:Contig30"
109280..170682
/note="assembly_name:Contig31"
clone_end:7"
BASE COUNT 43947 a 39732 c 39514 g 46884 t 605 others
ORIGIN

```

```

Query Match 31.4%; Score 896.2; DB 2; Length 170682;
Best Local Similarity 96.2%; Pred. No. 2.2e-151;
Matches 950; Conservative 0; Mismatches 34; Indels 4; Gaps 3;

```

```

QY 1 ATGCTTCTCTGGGATCCTTAACCTGGCTTTTCGGCGGCGAACCCTGGAGGCTCTGAG 60
Db 98460 ATGCTTCTCTGGGATCCTTAACCTGGCTTTTCGGCGGCGAACCCTGGAGGCTCTGAG 98519
QY 61 CCAG-AGCGGGAGGTA-GTCGTTCCCATCCGACTGGACCCGACATTAAACGCCGCCGCT 118
Db 98520 CCAGCAGCGGAGGTAAGCTGCTTCCCATTCGACTGGACCCGACATTAAACGCTTTCGCT 98579
QY 119 ACTACTGGCGGGTCCCGAGGAGCTCCGGGATCAGGAGCTCAATTTTCAGATCAGCAT 178
Db 98580 TTATCTGGCGGGTACC--GAGGACCCCGGGGTCCAGGNACTCAATTTTCAGATCAGCAT 98637
QY 179 TTTCAGGAGGACTTTTACCTACACCTGAGCGCGATGCTCAGTCTTGTGCTCCCGCTTCT 238
Db 98638 TTCAGGAGGACTTTTACCTACACCTGAGCGCGATGCTCAGTCTTGTGCTCCCGCTTCT 98697
QY 239 CCACTGAGCATCTGGCGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACCTTCGCAC 298
Db 98698 CCACTGAGCATCTGGCGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACCTTCGCAC 98757
QY 299 GCTGCTTCTTCTGGGAGCTGAACCGCGAGCGGAGCTCGTTCGCTGCTGTGAGCCTGT 358
Db 98758 GCTGCTTCTTCTGGGAGCTGAACCGCGAGCGGAGCTCGTTCGCTGCTGTGAGCCTGT 98817

```

QY	359	GGGGGGGGCTCCGGAGACGCTTTGGCTACCGAGGCGCGAGTATGTCTATTCATTCATTCAGCCCGCTGC	418	
Db	98818	GGGGGGGGCTCCGGAGACGCTTTGGCTACCGAGGCGCGAGTATGTCTATTCATTCATTCAGCCCGCTGC	98877	
QY	419	CCAATGCTAGCGCGCGGGCGAGCGACACAGCCAGGCGCGACACCTCTTCACGCGCC	478	
Db	98878	CCAATGCTAGCGCGCGGGCGAGCGACACAGCCAGGCGCGACACCTCTTCACGCGCC	98937	
QY	479	GGGGTGTTCGGGGCGGGCTTCGGAGAGCCACACTCTCGCTCGGGGGTGGCGCTTCGGGCT	538	
Db	98938	GGGGTGTTCGGGGCGGGCTTCGGAGAGCCACACTCTCGCTCGGGGGTGGCGCTTCGGGCT	98997	
QY	539	GGAAACCCCGCATCTACGGGCCCTTGAGACCTTACAAAGCCGCGGGCGGGCGCTTCGGGG	598	
Db	98998	GGAAACCCCGCATCTACGGGCCCTTGAGACCTTACAAAGCCGCGGGCGGGCGCTTCGGGG	99057	
QY	599	AGAGTCGTAGCGCGCGAGGCTCGGGCGCGCAAGCGTTTCGTCTATCCCGGGTACG	658	
Db	99058	AGAGTCGTAGCGCGCGAGGCTCGGGCGCGCAAGCGTTTCGTCTATCCCGGGTACG	99117	
QY	659	TGAGAGCGTGTGTGTCGGGACGAGTCAATGGTCAAGTCCACGGCGCGGACCTCGGAAC	718	
Db	99118	TGAGAGCGTGTGTGTCGGGACGAGTCAATGGTCAAGTCCACGGCGCGGACCTCGGAAC	99177	
QY	719	ATTATCTGCTGACGCTGTGGCAACGGCGCGGACTCTACCGGCATCCAGCATCTCTCA	778	
Db	99178	ATTATCTGCTGACGCTGTGGCAACGGCGCGGACTCTACCGGCATCCAGCATCTCTCA	99237	
QY	779	ACCCATCAACATCTGTGTGTCAGAGTCTCTTCTTAGAGATCGTGAATCCGGGGCCCA	838	
Db	99238	ACCCATCAACATCTGTGTGTCAGAGTCTCTTCTTAGAGATCGTGAATCCGGGGCCCA	99297	
QY	839	AGTCCACGGCAATCGGCCCTGAGCTGCGCAACTTCTGTCTGGCAGAGAGCTGA	898	
Db	99298	AGTCCACGGCAATCGGCCCTGAGCTGCGCAACTTCTGTCTGGCAGAGAGCTGA	99357	
QY	899	ACAAAGTGAGTGACAGACACCCCGAGTACTGGACACTGCGCATCTCTTACACGAGCAGG	958	
Db	99358	ACAAAGTGAGTGACAGACACCCCGAGTACTGGACACTGCGCATCTCTTACACGAGCAGG	99417	
QY	959	ACCTGTGTGGAGCCACCACCTGTGACAC	986	
Db	99418	TGAGTTGATCTGCGCTCACTTGTGCACCC	99445	
RESULT 7				
AC101990/c				
LOCUS	AC101990	182656 bp	DNA	linear
DEFINITION	Mus musculus clone RP24-371J2, WORKING DRAFT SEQUENCE, 15 unordered pieces.			HTG 21-AUG-2002
ACCESSION	AC101990			
VERSION	AC101990.2	GI:22381363		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.			
TITLE	Mus musculus, clone RP24-371J2			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 182656)			
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,			

```

* 1639 3397: contig of 1759 bp in length
* 3398 3497: gap of 100 bp
* 3498 4947: contig of 1450 bp in length
* 4948 5047: gap of 100 bp
* 5048 6893: contig of 1846 bp in length
* 6894 6993: gap of 100 bp
* 6994 10395: contig of 3402 bp in length
* 10396 10495: gap of 100 bp
* 10496 12584: contig of 2089 bp in length
* 12585 12684: gap of 100 bp
* 12685 14977: contig of 2293 bp in length
* 14978 15077: gap of 100 bp
* 15078 18181: contig of 3104 bp in length
* 18182 18281: gap of 100 bp
* 18282 22654: contig of 4373 bp in length
* 22655 22754: gap of 100 bp
* 22755 31263: contig of 8507 bp in length
* 31262 31361: gap of 100 bp
* 31362 43587: contig of 12226 bp in length
* 43588 43687: gap of 100 bp
* 43688 57612: contig of 13925 bp in length
* 57613 57712: gap of 100 bp
* 57713 77467: contig of 19755 bp in length
* 77468 77567: gap of 100 bp
* 77568 108724: contig of 31157 bp in length
* 108725 108824: gap of 100 bp
* 108825 182656: contig of 73832 bp in length.

```

FEATURES

```

source
  1. 182656
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone_lib="RP24-371J2"
    /clone_lib="RPCI-24 Male Mouse BAC"
  1. 361
    /note="assembly_fragment"
  462. 1538
    /note="assembly_fragment"
  1639. 3397
    /note="assembly_fragment"
  3498. 4947
    /note="assembly_fragment"
  5048. 6893
    /note="assembly_fragment"
  6994. 10395
    /note="assembly_fragment"
  10496. 12584
    /note="assembly_fragment"
  12685. 14977
    /note="assembly_fragment"
  15078. 18181
    /note="assembly_fragment"
  18282. 22654
    /note="assembly_fragment"
  22755. 31261
    /note="assembly_fragment"
  31362. 43587
    /note="assembly_fragment"
  43688. 57612
    /note="assembly_fragment"
  57713. 77467
    /note="assembly_fragment"
  77568. 108724
    /note="assembly_fragment"
  108825. 182656
    /note="assembly_fragment"
  1500 others
BASE COUNT 51468 a 41160 c 39770 g 48758 t
ORIGIN

```

```

Query Match 27.4%; Score 783; DB 2; Length 182656;
Best Local Similarity 87.3%; Pred. No. 4.5e-131;
Matches 858; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 1 ATGCTTCTGTGGCATCTCAACCTGGCTTCGCCGGCGGAACCGCTGAGGCTCTGAG 60

```

```

Db 39782 ATGCTTCTGTGGCATCTCAATCTGGCTCTGCCCTGGGAGCTGCTGGCAGCTCCGAG 39723
QY 61 CCAGAGCGGGAGGTAGTCGTTCCATCCGACTGGACCGGACATTAACGCGCGGCTAC 120
Db 39722 CCAGAGTGGAGGTGCTGTTCCATCCGAGCGGACCGGACATCAATGCGCGGCTAC 39663
QY 121 TACTGGCGGGTCCGAGGAGTCCGGGATCAGGAGTCAATTTTTCAGATCAGAGATTT 180
Db 39662 TACAGAGGGGTACGGAGGACTCCGGGATCAGGGTCTCATTTTTCAGATCAGAGTCTT 39603
QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCCTGCTCCGCTCTCTCC 240
Db 39602 CAGCAGGACTTTTATCTACACCTGACGCGGATGCTCAGTTCCTGCTCCGCTCTCTCC 39543
QY 241 ACTGAGCATCTGGCGCTCCCTCCAGGGCTCACCGGGGCTCTTCAGACTTCGCGACGC 300
Db 39542 ACTGAGTATCTAGGTGCTCCGCTGCAGAGGCTCCTAGGAGCTCTAGACCTTGCACGC 39483
QY 301 TGCTTCTATTCTGGGACGTGAACGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 39482 TGCTTCTATTCTGGGATGATGAACGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 39423
QY 361 GGGGGCTCCGCGGAGCTTTGGCTACCGAGCGCGGAGTATGCTATAGCCGCTGCTGCTG 420
Db 39422 GGGGGTCTCCGCGGAGCTTTGGCTACCGAGGTGCGGAGTATGCTATAGCCCTCTGCCC 39363
QY 421 AATGCTAGCGCGCGGCGCGAGCGCAACAGCGAGGGCGCACACTTCTCCAGCGCGG 480
Db 39362 AACCCAGCGCGCGAGAGCGCGAGCTCACAGCCAGGGCGCACACTTCTCCAGCGCGG 39303
QY 481 GGTGTTCCGGCGGCGCTTCCGAGAGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 39302 GGTGCTCTCTAGGGCTTCCGGAGAGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 39243
QY 541 AACCCGCGCATCTACGGCGCTTGGACCTTACAAGCGCGGCGGCGGCTTCCGGCGGAG 600
Db 39242 AACCCGCGCATCTGAGGGCTCTGAGACCTTATAAGCCAGCGGCGGCGGCGGAG 39183
QY 601 AGTCGTAGCGCGGCGAGGTCTGGCGCGCGCAAGCGTTTCTGCTATCCCGGCTACGTG 660
Db 39182 AGCCACAACCGCGCGAGGTCTGGCGCGCGCAAGCGTCTGCTATACAGCGTACGTG 39123
QY 661 GAGACGCTGCTGCTGCGGAGGAGTCAATGGTCAAGTTCACGGCGCGGAGCTCCCAAC 720
Db 39122 GAGACACTGCTGCTGCGGAGGAGTCAATGGTCAAGTTCACGGCGCGGAGTTGGAAC 39063
QY 721 TATCTCTGACGCTGCTGCGGAGGAGTCAATGGTCAAGTTCACGGCGCGGAGTTGGAAC 780
Db 39062 TATCTCTGACGCTGCTGCGGAGGAGTCAATGGTCAAGTTCACGGCGCGGAGTTGGAAC 39003
QY 781 CCCATCAACATGCTGCTGCTGCGGAGGAGTCAATGGTCAAGTTCACGGCGCGGAGTTGGAAC 840
Db 39002 CCTATCAACATGCTGCTGCTGCGGAGGAGTCAATGGTCAAGTTCACGGCGCGGAGTTGGAAC 38943
QY 841 GTACCGGCAATGCGGCGCTGAGCTGCGCAACTTCTGCTGCTGCGGAGGAGTCAAGTCAAC 900
Db 38942 GTACAGGCAACGCGGCGCTGACTCTGCGCAACTTCTGCTGCTGCGGAGGAGTCAAGTCAAC 38883
QY 901 AAAGTGAAGTGAACAGCACCGGAGTCTGGGAGTCTGCGCAACTTCTGCTGCTGCGGAGGAGTCAAGTCAAC 960
Db 38882 AAAGTGAAGTGAACAGCACCGGAGTCTGGGAGTCTGGGAGTCTGGGAGTCTGGGAGTCTGGGAGT 38823
QY 961 CTGTGTGGAGGACCACTGCTGA 983
Db 38822 AGTTCATCTGTCACTCTTTGA 38800

```

RESULT 8

AC126507

LOCUS

DEFINITION

Rattus norvegicus clone CH230-254N12, *** SEQUENCING IN PROGRESS

***, 49 unordered pieces.

AC126507

178764 bp

DNA

linear

HTG 24-JUL-2002

AC126507
 VERSION AC126507.1 GI:21700463
 KEYWORDS HTG: HTGS_PHASE1
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 1 (bases 1 to 178764)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaraturge,H.C., Are,J.R., Ayelle,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhaý,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,F., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuben,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

Direct Submission
 Unpublished
 2 (bases 1 to 178764)
 Worley,K.C.
 Direct Submission
 Submitted (06-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 3 (bases 1 to 178764)
 Worley,K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GZFU
 Center clone name: CH230-254N12
 ----- Summary Statistics
 Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 124573 bases at least Q40
 Consensus quality: 131850 bases at least Q30
 Consensus quality: 137700 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 49 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 1037: contig of 1037 bp in length
 1137: gap of unknown length
 1138: contig of 1113 bp in length
 2251: gap of unknown length
 2351: contig of 1229 bp in length
 3580: gap of unknown length
 3680: contig of 1195 bp in length
 4875: gap of unknown length
 4874: contig of 1195 bp in length
 4875: gap of unknown length
 4975: contig of 1551 bp in length
 6526: gap of unknown length
 6626: contig of 1280 bp in length
 7906: gap of unknown length
 8005: contig of 1719 bp in length
 9725: gap of unknown length
 9825: contig of 1618 bp in length
 11442: gap of unknown length
 11443: contig of 1358 bp in length
 11543: gap of unknown length
 12901: contig of 1150 bp in length
 13001: contig of 1150 bp in length
 14151: gap of unknown length
 14250: contig of 1315 bp in length
 14565: gap of unknown length
 15566: contig of 1810 bp in length
 15666: gap of unknown length
 17476: contig of 1115 bp in length
 17575: gap of unknown length
 18691: contig of 1479 bp in length
 18790: gap of unknown length
 20269: contig of 1558 bp in length
 20367: gap of unknown length
 20370: contig of 1558 bp in length
 21927: gap of unknown length
 21928: contig of 1971 bp in length
 23998: gap of unknown length
 24098: contig of 1581 bp in length
 25679: gap of unknown length
 25779: gap of unknown length
 25780: contig of 1649 bp in length
 27429: gap of unknown length
 27529: contig of 2149 bp in length
 29677: gap of unknown length
 29778: contig of 1100 bp in length
 30877: gap of unknown length
 30977: gap of unknown length
 32831: contig of 1854 bp in length
 32832: gap of unknown length
 32931: contig of 2488 bp in length
 35419: gap of unknown length
 35520: contig of 1874 bp in length
 37393: gap of unknown length
 37493: contig of 2342 bp in length
 39835: gap of unknown length
 39935: contig of 2854 bp in length
 42789: gap of unknown length
 42889: contig of 3757 bp in length
 46546: gap of unknown length
 46746: contig of 2442 bp in length
 49188: gap of unknown length
 49288: contig of 4517 bp in length
 53605: gap of unknown length
 53905: contig of 2297 bp in length
 53906: contig of 2297 bp in length

COMMENT		FEATURES		source	
Tel:81-45-503-9111, Fax:81-45-503-9170)					
On Mar 7, 2002 this sequence version replaced gt:13488920.					
Location/Qualifiers					
1..28000					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
/chromosome="11"					
/map="11q"					
/clone="RP11-211H5"					
BASE COUNT 6938 a 6083 c 6679 g 8300 t					
ORIGIN					
Query Match 27.2%; Score 777.2; DB 9; Length 28000;					
Best Local Similarity 98.4%; Pred. No. 7.2e-130;					
Matches 785; Conservative 0; Mismatches 13; Indels 0; Gaps 0;					
QY	2056	ACTGTGACTCTTCCACCAAGCCCATGTCATGGCTACAAATTCGTGTGGCGCATCCCGCAGGC	2115		
DB	3450	ACGGCCCCCTTCCCGCGCAGGCATGGCTACAAATTCGTGTGGCGCATCCCGCAGGC	3509		
QY	2116	GCCTCAAGCATCGACATCCGCCAGCGCGGTTACAAAGGGCTGATCGGGGATGACAACTAC	2175		
DB	3510	GCCTCAAGCATCGACATCCGCCAGCGCGGTTACAAAGGGCTGATCGGGGATGACAACTAC	3569		
QY	2176	CTGGCTCTGAAGAACAACGCCAAGCAAGTACCTGCTCAAGGGCATTTCTGGTGGCGG	2235		
DB	3570	CTGGCTCTGAAGAACAACGCCAAGCAAGTACCTGCTCAAGGGCATTTCTGGTGGCGG	3629		
QY	2236	GTGGAGCGGGACCTGGTGTGAAGGGCAGTGTCTGTGGGGTACAGCGGCACGGGCACAGCG	2295		
DB	3630	GTGGAGCGGGACCTGGTGTGAAGGGCAGTGTCTGTGGGGTACAGCGGCACGGGCACAGCG	3689		
QY	2296	GTGGAGAGCTGCAGAGCTTCCCGGCCCATCTCTGGAGCGCTGACCGTGGAGGTCCTCTCC	2355		
DB	3690	GTGGAGAGCTGCAGAGCTTCCCGGCCCATCTCTGGAGCGCTGACCGTGGAGGTCCTCTCC	3749		
QY	2356	GTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTGCCCCAAGAGCCTCGG	2415		
DB	3750	GTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTGCCCCAAGAGCCTCGG	3809		
QY	2416	GAGGACAAGTCTCTCATCCCAAGNACCCCGGGAGCCCTGTGTTGCACACACAGCGTC	2475		
DB	3810	GAGGACAAGTCTCTCATCCCAAGNACCCCGGGAGCCCTGTGTTGCACACACAGCGTC	3869		
QY	2476	CTCAGCCTCTCCCAACAGGTGGAGCAGCGGACAGCGCCCTCTGACACCTGGGTGGCT	2535		
DB	3870	CTCAGCCTCTCCCAACAGGTGGAGCAGCGGACAGCGCCCTCTGACACCTGGGTGGCT	3929		
QY	2536	GGCAGCTGGGGGCCGTGCTCCGGAGCTGCGGCAGTGGCCTGCAGAGACGGGGCGGTGGAC	2595		
DB	3930	GGCAGCTGGGGGCCGTGCTCCGGAGCTGCGGCAGTGGCCTGCAGAGACGGGGCGGTGGAC	3989		
QY	2596	TGTCGGGGCTCCCGGGCAGCGCAGGTCCTGCTGTGATGCAGCCCATCGGCCCGGTG	2655		
DB	3990	TGTCGGGGCTCCCGGGCAGCGCAGGTCCTGCTGTGATGCAGCCCATCGGCCCGGTG	4049		
QY	2656	GAGACACAAGCCTCGCGGGAGCCCTGCCCCACTCTGGGAGCTCAGCGCTGTGTCACCCCTG	2715		
DB	4050	GAGACACAAGCCTCGCGGGAGCCCTGCCCCACTCTGGGAGCTCAGCGCTGTGTCACCCCTG	4109		
QY	2716	TCCAAGAGCTGGCGCGGGATTTACAGAGCCCTCACTCAAGTGTGTGGCCACGAGGC	2775		
DB	4110	TCCAAGAGCTGGCGCGGGATTTACAGAGCCCTCACTCAAGTGTGTGGCCACGAGGC	4169		
QY	2776	CGGCTGCTGGCCGGGACCAAGTGCAACTTGCACCCCAAGCCCCCAGGAGCTGGACTTCTGC	2835		
DB	4170	CGGCTGCTGGCCGGGACCAAGTGCAACTTGCACCCCAAGCCCCCAGGAGCTGGACTTCTGC	4229		
QY	2836	GTCTGTAGGCGGTGCTGA 2853			
DB	4230	GTCTGTAGGCGGTGCTGA 4247			

QY 751 CGACTCTACGGCCATCCACGATCTCTCAACCCATCAATCGTGTGTCAAGTGCTG 810
DB 754 CGAATCTACAAGCACCACGATCAAGATTCATCACTCATGTGTGTAAGTGCTG 813
QY 811 CTCTTTAGAGATGTGACTCCGGGGCCCAAGGTACACGGCAATGCGGCCCTGACGCTGGCC 870
DB 814 ATCGTAGAAGTGAATAATGGGGCCAGAGGTGTCCGAAATGGGGCTTACACTGGGT 873
QY 871 AACTCTGTGCTGGCAGAGAGCTGACAAAGTGAAGTGAAGCAACCCCGAGTACTGG 930
DB 874 AACTCTGTGCAACTGGCAGCGGGTTTCAACCCAGCCGAGCGCCACCCAGAGCACTAC 933
QY 931 GACACTGCCATCTCTTCAACGAGGAGGACTGTGTGG---AGCCACCACCTTGACACC 987
DB 934 GACAGGCCATCTGCTCAACGAGACAACTTCTGTGGCAGGAGGGGTGTGTGACACC 993
QY 988 CTGGGANTGCTGATGTGGGTAGCATGTGTGACCCCAAGAGAGCTGCTCTCATTTGAG 1047
DB 994 CTGGGTGGCAGACATCGGGGACCACTTGTGACCCCAAGAAAGCTGCTCCGTGATCGAG 1053
QY 1048 GAGGATGGGCTTCCATCAGCTTCAACACTGCCACGAGCTGGCCACCGTGTCAACATG 1107
DB 1054 GATGAGGGGCTCAGGGGGCCACACCTCGGCCCATGAACAGGACAGCTCTCAGCATG 1113
QY 1108 CCCCATGACAAATGTGAAGTCTGTGAGGAGGTGTTTGGGAAGCTCCGAGCCCAACACATG 1167
DB 1114 CCCACGAGACTCCAAAGCCCTCACACGCGCTCTCGGGCCCATGGCAAGCACACGCTG 1173
QY 1168 ATGTCCCGACCTCATTCAGATGACGGTGGCCAAACCCCTGGTCAGCCTGCGAGTGGTCC 1227
DB 1174 ATGGCAGCGTGTGTCTGCTCAACCTGAACAGAGCTGCCCTGTGCTCCCGTCCAGCGCATG 1233
QY 1228 ATCATCACGACTTCTGGACACGGCGAGTGTGCTCTGTGACCAACCCAGCAAG 1287
DB 1234 TATCTCACAGCTTCTGGACGGGGGACGAGAGTGTCTCTGATGCCCTGCTGCTGG 1293
QY 1288 CCATCTCTCCCTGGCCGAGTGTGCGGGG-----GCCAGTACACCTTGAGCCAGCAG 1341
DB 1294 GCCCTGCCCTCCACAGGCTCCCGGGCGCATGGCCCTGTACCAAGTGGACAGCAG 1353
QY 1342 TGGAGTGTGCTTTGGCGTGGGCTCAAGCCCTGTCTTACA-----TGCAGTAC 1392
DB 1354 TGCAGGAGATCTTTGGGGCGGATTTCCGGCCACTGCCCAACACCTCTGCTCAGGACGTC 1413
QY 1393 TGCACCAAGCTGTGGTG---CACCGGAGGCCAAGGACAGATGCTGTGCCAGACCCGC 1449
DB 1414 TGCGCCAGCTTTGGTCCACACTGATGGGGCTGAGCCCTGTGCCACAGAGAATGGC 1473
QY 1450 CACTTCCCTGGCCGATGGCACAGCTGTGGCGAGGGCAAGCTCTGCCCTCAAGGGGCC 1509
DB 1474 AGCCTGCCCTGGGTGACGCGCACGCCCTGGGGCCCTGGGCACTCTGCTCAGAAGGACG 1533
QY 1510 TG-----CGTGGAGACACACACTCAACAGCAGAGGTGATGTTCTTGGGCCAA 1563
DB 1534 TGTCTACTGAGGAGGAAGTGGAGAGGCCCAAGCCCTGTGTAGATGGAGCTGGGCACCG 1593
QY 1564 TGGGATCCCTATGGCCCTGCTCGGCACATGTGTGGGGGCGCTGCGAGCTGGCCAGGAG 1623
DB 1594 TGGGGACCTTGGGAGATGTTCTGGACCTGTGGAGAGAGTACAGTTTACACACCT 1653
QY 1624 CAGTGCACCAACCCCTGCGCAACGCGGGGCAAGTACTGCGAGGAGTGGAGGTGAA 1683
DB 1654 GAGTGCAGAGGACCCGACGCTCAGATGGAGGAAGATACTGCTCGGTGGGTGGAGAGCAAG 1713
QY 1684 TACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCCGGAAGAGCTTCCGGGAG 1743
DB 1714 TACCGATCTGCGACGAGGAGATGGCCCC-----CCTACGGGAAAGCTTACGGGAG 1767
QY 1744 GAGCAGTGTGAGGCTTTCAACGGGTACACACGACAGCAACCCGCTCACTCTCGCCGTG 1803
DB 1768 CAGCAGTGTGAGAGTATATGCTACAAATTACACTGACATGGACGGGAATCT---CCTG 1824

QY 1804 GCATGGGTGCCAAGTACTCCGGGGTGTCTCCCGGACAAAGTGCAGCTCATCTGCGGA 1863
DB 1825 CAGTGGGTCCCAAGTACTGCTGGGGTGTCCCGGACCGCTGCAAGTGTGTTCTGCGGA 1884
QY 1864 GCCAATGSCACTGCTACTTCTATGCTGTCGACCCCAAGTGTGGACGCGCCTGTGC 1923
DB 1885 GCCCGGGGAGGAGCGAGTTCAAAGTGTTCGAGGCCAAGTGTATGATGACCCCTGTGT 1944
QY 1924 TCTCCTGACTCCACCTCCGCTCTGTCTCAAGCAAGTGCATCAAGGCTGCTGTGATGG 1983
DB 1945 GGGCCAGAACACATGCGCATCTGTGTCGTCGTCAGTGTGTCAAGGCCGCTGTGACCAT 2004
QY 1984 AACCTGGGCTCCAAAGAGATTCGACAAGTGTGGGGTGTGTGGGGGAGACATAAGC 2043
DB 2005 GTGGTGAAGTCTCGCTCGGAAGCTGGACAAATGCGGGGTGTGTGGGGCAAGGCAATCC 2064
QY 2044 TGCAGAGGTGACTGTGACTCTTCCACCAAGCCATGCTACATTTTCGTGGTGGCC 2103
DB 2065 TGCAGGAGGTCTCCGGTCCCTCACCCACCAANTATGCTACATGACATTTGTCACC 2124
QY 2104 ATCCCGCAGCGCTCAAGCATCGACATCGCGAGCGGTTTACAAAGGCTGATCGGG 2163
DB 2125 ATCCAGCTGTGCGCACTAATATTGAGTGAAGCAGCGGAGCCACCGGGTGTGCAAGC 2184
QY 2164 GATGACAACTACTGCTCTGAAGAACAGCAAGGCAAGTACCTGCTCAAGCGCATTTTC 2223
DB 2185 GATGGAACTACTGCGGCTGAAGACGGCTGATGGGAGTACCTGCTCAAGCGCAACCTG 2244
QY 2224 GTGGTGTGCGGGTGGAGCGGAGCTGTGGTGAAGGCGAGTCTGTGCGGTACAGCGGC 2283
DB 2245 GCATCTCTGCCATGAGCAGACATCTTGTGAAGGGGACCATCTTGAAGTACAGCGGC 2304
QY 2284 ACGGGACAGCGGTGGAGAGCTCTCAGGCTTCCCGGCCCATCTGTGGAGCGCTGACCGTG 2343
DB 2305 TCCATGCGCACCTTGGAGCGCTCTCAGAGCTTCCGGGCTTGCAGAGACCTCTGACAGTG 2364
QY 2344 GAGTGTCTCTCCGT---GGGGAAGATGACACCGCCCGGGTCCCTACTCTCTTCTATCIG 2400
DB 2365 CAGTCTCTGACAGTCCCTGCGGAGGTCTTCCCCCAAAAGTCAANTACACCTTCTTGT 2424
QY 2401 CCCAAGA 2408
DB 2425 CCTAATGA 2432
RESULT 12
E58655 Novel metallic protease.
LOCUS E58655
DEFINITION E58655
ACCESSION E58655.1 GI:18629877
VERSION JP 2001017183-A/3.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2670)
AUTHORS Yamaji,N., Nishimura,K. and Sasamata,M.
TITLE Novel metallic protease
JOURNAL Patent: JP 2001017183-A 3 23-JAN-2001;
YAMANOUCHI PHARMACEUT CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2001017183-A/3
PD 23-JAN-2001
PF 09-JUL-1999 JP 1999196584
PR NOBORU YAMAJI,KOICHI NISHIMURA,MIHO SASAMATA
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC
C12N9/50,C12Q1/37,
CC C12N15/00,C12N5/00
CH Key Location/Qualifiers
FT source 1..2670

FEATURES		FT	Location/Qualifiers		/organism='Homo sapiens (human)'	
source			1..2670			
BASE COUNT			511 a 853 c 867 g 439 t			
ORIGIN						
Query Match 23.3%; Score 664.4; DB 6; Length 2670;						
Best Local Similarity 59.2%; Pred. No. 2e-109;						
Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;						
QY	157	CTCATTTTTCAGATCACAGCAATTCAGGAGGACTTTTACCTACACCTGATGCCGATGCT	216			
Db	157	CTCGGGCTCCACTTCGGCTTCGGCAAGGGCTTCGTGCTCGCGCTTGGTCCCGAGAC	216			
QY	217	CAGTTCTTGGCTCCGCTTCCTCACTAGCATCTGGCGTCCCTCTCCAGAGGCTGCT	276			
Db	217	AGCTTCTTGGCGGCGAGTTCAAGATCGAGCGCTTCGGGGGCTCGGGCGGCG---ACC	273			
QY	277	GGGGCTCTTCAGACCTCGAGCTGCTTCTATTCTGGGGAGCTGACGCTCGAGCCGAC	336			
Db	274	GGGGCGAGCGGGGCTCGCGGCTGCTTCTTCTCCGACCGTGAATGGGGAGCCCGAG	333			
QY	337	TGCTTCGCTGCTGTAGCCTGTGGGGGGGCTCCGCGAGGCTTGTGGTATCGAGGCGCC	396			
Db	334	TGCTTGGCGGGGTCAGCCTGTGCGCGGGGTGAGCGGCTCTTCTGCTGGAGCGGAG	393			
QY	397	GAGTATGTCTATTAGCCCTGCCAATGCTAGCGCGCGGGGCTTCGGAGATCCCACTCT	456			
Db	394	GAGTTCACCATCCAGCCGAGCGCGGGGCTCCCTGCGCTCAGCCGCTGAGCGCTGAG	453			
QY	457	GGCGCACACCTTCTCCAGCGCGGGGTTTCCGGGCGGGCTTCGGAGATCCCACTCT	516			
Db	454	CGCTGGGCTCCCGCGGAGCGCGCGCTCCCGGAGGACCCGAGTGGGAGTGGAGAGC	513			
QY	517	CGCTCGGGGTG-CCCTGGGCTGGAACCCCGGCATCTACGGGCGCTTGGACCTTACAA	575			
Db	514	GGAGAGGTTCAAGCGAGGAGAGGAGGACCAACGAGGAGGAGGAGGAGGAGGACAA	573			
QY	576	GCCCGCGCGGGCTTCGGGGAGAGTCGTGAGCGGGCGAGG-----TCTGGGCGGCC	630			
Db	574	GAAGAGGAGGAGAGGCGCTAGCGCGCGCCAGCGCCCTTGGGGCCACAGTAGGACC	633			
QY	631	AAGCGTTTCGTCTATTCGCCGCTAGTGGAGAGCTGTGTGGTGGGAGGAGTCAATG	690			
Db	634	AAGCGGTTTGTCTGAGCGCGCTTCTGTAAGAGCTGCTGTGGTGGCGGATGCTCCATG	693			
QY	691	GTCAAGTTTCCAGCGCGGACCTGGAACATTTCTGCTGAGCTGTGGGTAACGGCGCG	750			
Db	694	GCTGCTTCTACGGGGCGGCTTCAGAAACACATCTGAGGTTTAATGCTGTGGCAGCC	753			
QY	751	CGACTCTACCGCCATCCAGCATCTCAACCCCATCAACATCTGTGGTCAAGTCTGTG	810			
Db	754	CGAATCTACAGCACCCAGCATCAAGATTTCAATCACTGATGTGGTAAAGTCTGTG	813			
QY	811	CTTCTTAGAGATGCTGACTCCGGGCCCAAGTCAACCGCAATGGGCCCTGACGCTGCGC	870			
Db	814	ATCGTAGAAGATGAAATGGGGCCAGAGGTGTCCGACAAATGGGGGCTTACACTGGT	873			
QY	871	AACCTTCTGCTGGCAGAGAGAGCTCAACAAGTGTGACAGACACCGGAGTACTGG	930			
Db	874	AACCTTCTGCACTGGGAGCGGGCTTTCACACGAGCCGACGACCCACCGAGACATAC	933			
QY	931	GACACTGCCATCTCTTACAGCAGGAGGACTGTGTGG---AGCCACCACTGTGACACC	987			
Db	934	GACAGGCCATCTGCTCACACAGACAGAACTTCTGTGGGAGGAGGGGCTGTGTGACCC	993			
QY	988	CTGGGATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGCTGTTCATTGAG	1047			
Db	994	CTGGGTGTGCAGACATCGGGAGCAATTTGTGACCCCAACAAAGCTGCTCTGTGATCGAG	1053			
QY	1048	GACGATGGGCTTCATCAGCCTTCAACACTGCCACAGCTGGGCGGCTTACAAAGGCTGATCGG	1107			

Db 2125 ATCCAGCTGGTCCCACTAATATGAGTGAAGCAGCGGACCCCGGCTGTGCAGAAC 2184

Qy 2164 GATGACAACTACCTGCTCTGAAGAACAGCCAGGCAAGTACTGCTCAACGGGCATTTC 2223

Db 2185 GATGGGAACCTACCTGGCTCTGAAGAGCGGTGATGGCGAGTACTGCTCAACGGCACTTG 2244

Qy 2224 GTGGTGTGCGGCTGGAGGGGACCTGGTGGTGAAGGGCAGTCTGCTCGGCTACAGCGGC 2283

Db 2245 GCCATCTCTGCCATAGCAGGACATCTTGGTGAAGGGACCATCTCTGAAGTACAGCGGC 2304

Qy 2284 ACGGGACACGGGTGAGAGCCTGAGGCTTCGCCGCCCATCTCTGAGGCGGTGACCGTG 2343

Db 2305 TCCATGCCACCTGGAGCGCTGACAGCTTCGGGCCCTTGCACAGAGCTCTGACAGTG 2364

Qy 2344 GAGTCTCTCTCCGT---GGGGAAGATGACACGCCGCCGGTCCGCTACTCTCTATCTG 2400

Db 2365 CAGTCTCTGACAGTCCCTGGCGAGGTCTTCCGCCCAAAAGTCAATATACACTCTCTTGT 2424

Qy 2401 CCCAAAGA 2408

Db 2425 CCTAATGA 2432

RESULT 13

AF060153

LOCUS

DEFINITION Homo sapiens METH2 protein (METH2) mRNA, complete cds.

ACCESSION AF060153

VERSION AF060153.1 GI:5725507

KEYWORDS

SOURCE

ORGANISM

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3711)

Vazquez,F., Hastings,G., Ortega,M.A., Lane,T.F., Oikemus,S., Lombardo,M., and Iruela-Arispe,M.L.

METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity

J. Biol. Chem. 274 (33), 23349-23357 (1999)

9367466

10438512

2 (bases 1 to 3711)

Vazquez,F., Hastings,G., Ortega,M.-A., Lane,T.F., Lombardo,M., Oikemus,S. and Iruela-Arispe,M.L.

Direct Submission

Submitted (16-APR-1998) Pathology, Beth Israel Deaconess Medical Center, 99, Brookline Avenue, Boston, MA 02214, USA

Location/Qualifiers

1. .3711

/organism="Homo sapiens"

/db_xref="taxon:9606"

1. .3711

/gene="METH2"

706..3378

/gene="METH2"

/note="extracellular protein with metalloprotease and TSP domains"

/codon_start=1

/product="METH2 protein"

/protein_id="A048081.1"

/db_xref="GI:5725508"

/translation="MPPAAPRWLFLLLLLLLPLARGAPARPAAGQASLVVP

TRLPAGAGELALHLSAFKGFVLRAPDDSFAPFKIERLGGSGRATGERLGRGCF

FSGTVNGEPLASLRCGLSGFLDGEFTIQPGAGGSLAQPHRLQWRGAPGAR

PLPRGEWETEGORQERGHODESEESQEEAEGASEPPPLPGATSRTRKRVSE

ARVETLLVADASMAAFYADLONHILLTMSVAARYKHPISIKNSINLMVVKLVIED

EKGPEVSDNGGLTLNFCNWRFRNPQSDRPEHYDTALLTRONFCQEGLCDTLG

VADIGTIDPNKSCSVIEDGLQAHLAHLGHLVLSMHDSDSKPTRLFGPMGRHHV

MAPFLVHLNLTLPSPCSAMYLTELLDGGHGDCLLDAPGAALPPTLFGRLMALYOLD

QQCRQLQFDFRCPRTSAQVCAQLWCHTKNGSLPWADGTGCPGPHLC

SESGCLPEEERPKVPDVGWAPWPGEGECRTCCGGGVQFVSHRECKDPENGRYC

LGRRAVQSCHTECPDGKSPFEQCEKYNAVNTDMDGNLLQWPKVAGYSPDRDC

KLFRCARGRSEFKVFEAKVIDGTLCGPETLAIICVRQCQKAGCDHVDSPRKLDKGV

CGKGNCSRKVSGLTPTNYGNDIVITIPAGATNIDVOKRSPGVNDGNLALKTAD

CGKGNLAIISAEIQDILVKYSGSIATLRFQSPRLPTVOLLTVPGEV

FPKVKYTFPVPNDVDFSMOSSKERATNLIQPLHAQWLVGDWSECSTCGAGWRR

TVFCRPSGOASATCNKALPKPAKESQICPL"

BASE COUNT 733 a 1126 c 1201 g 646 t 5 others

ORIGIN

Query Match 23.3%; Score 664.4; DB 9; Length 3711;

Best Local Similarity 59.2%; Pred. No. 1.9e-109;

Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;

Qy 157 CTCATTTTTCAGATCAGACATTTTCAGAGGACCTTTTACCTACACCTGAGCCCGATGCT 216

Db 865 CTCGGCTCCACCTGTCGCCCTTCGGCAAGGCTTCGTGTTCGCCCTGGCGCCGACAC 924

Qy 217 CAGTCTTGGCTCCGCCCTTCTCCACTGAGCATCTGGCGTCCCTCCAGGGCTCAC 276

Db 925 AGCTTCTGGCGCCCGAGTTCAGATCGAGCGCTCGGGGCTCCGGCCGGCG--ACC 981

Qy 277 GGGGCTCTTCAGACCTCGACGCTGCTTCTATTCTGGGACGTGAACGCCGACGCCGAC 336

Db 982 GGGGCGAGCGGGCTGCGCGCTGTTTTTTCGGCACCGTGAATGGGAGCCCCAG 1041

Qy 337 TCGTTCGCTGTGAGCCTGTGCGGGGGCTCCGGGAGCCTTTGGCTACCGAGCGCC 396

Db 1042 TCGCTGGCGGGTTCAGCTGTGCGGGGCTGAGCGCTCTCTTCCTGCTGAGCGGGAG 1101

Qy 397 GAGTATGTCATTAGCCCGCTGCCAATGCTAGCGCGCGCGGCGCAGCGCAACAGCAG 456

Db 1102 GAGTTCACCATCCAGCCAGCGCGGGGGCTCCCTGGCTCAGCGCACCGCTGCAG 1161

Qy 457 GGGCACACCTTCTCAGCGCGGGGGTGTTCGGCGGGGCTTCGGGAGACCCACCTCT 516

Db 1162 CGCTGGGTCCCGCGGAGCGCCCTCCCGGAGGACCCGAGTGGGAGTGGAGACG 1221

Qy 517 CGCTCGGGGTG-GCTCGGCTGAACCCCGCATCTACGGCGCTTGACCTTACAA 575

Db 1222 GGAGAGGTCAGAGGACGAGAGGAGGACCCAGGAGGACGAGGAGGAGGAGAGUCAA 1281

Qy 576 GCGCGCGGGCGGGCTTCGGGGAGAGTCGTAGCGCGCGCAGG-----TCTGGCGGCC 630

Db 1282 GAAGAGGAGGAGAGGCGCTAGCGCGCGCCACCGCCCTGGGGCCACGAGTAGGACC 1341

Qy 631 AAGCTTTGCTGTCTATCCCGGGTACGTGAGAGCGCTGTGTGTCGGGACGAGTCAATG 690

Db 1342 AAGCGGTTTGTCTGTGAGCGCGCTTCGTGGAGAGCTGCTGGTGGCGGATGCGTCCAIG 1401

Qy 691 GTCAAGTTCACGCGCGGACCTGGAACATATCTGTGACGCTGCTGGCAACGCGCGCG 750

Db 1402 GTGCTCTTCTACGGGGCGGACCTGCAGAACACATCTGACGTTAATCTCTGTGGACCC 1461

Qy 751 CGACTCTACCGCATCCAGCATCTCAACCCCATCAACATCTGTTGGTCAAGTGTG 810

Db 1462 CGAATCTACAAGCACCCAGCATCAAGAAATTCATCAACCTGATGGTGTAAAAGTGTG 1521

Qy 811 CTTCTAGAGATCTGACTCCCGGGCCCAAGGTCACCGCAATGCGGCTTACGCTGCGC 870

Db 1522 ATCTGTAGAGATGAATAATGGGGCCCGAGGTTGTCGGAATATGGGGGCTTACACTGCT 1581

Qy 871 AACTTCTGTGCTGGCAGAAAGCTGAACAAAGTGAAGTGAACACCCGAGTACTGG 930

Db 1582 AACTTCTGAACTGGCAGGGGGTTCACACGACCCAGCGACCCAGAGCACTAC 1641

Qy 931 GACACTGCCATCTTTCACAGGAGGACCTGTGTGG---AGCCACCACTGTGACACC 987

Db 1642 GACACGGCCATCTGTCTACCAAGACAACTTCTGTGGCAGGAGGGGCTGTGTGACACC 1701

Qy 988 CTGGGCTGCTGATGTGGTACCATGTGTGACCCCAAGAGAGTGTCTGTCTCATTTAG 1047

Db 1702 CTGGGTGTGGCAGACATCGGGGACCATTTGTGACCCCAACAAAGTGTCTCGTGTGAT 1761

Qy 1048 GACGATGGCTTCCATCATGACCTTACCCTGCCCCACGAGCTGGGCCACGTGTTCACATG 1107

Db 1762 GATGAGGCTCCAGGCGCCACACCTTGGCCCTGAACTAGGCGACGCTCAGCATG 1821
QY 1108 CCCATGACAATGTGAAGTCTGTGAGGAGGTGTTTGGGAAGCTCCGAGCGCAACCAATG 1167
Db 1822 CCCACGAGCACTCCAAAGCCCTGACACCGCTTTCGGGCCCATGGGCAACACCAATG 1881
QY 1168 ATGTCCCGACCCCTCATCAGATCAGCGTGCACCCCTGTGAGCCCTGAGTGCAGTGC 1227
Db 1882 ATGGCACCGCTTGTGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1941
QY 1228 ATCATCACCGACTTCTGACAGCGGCGACGCTGACTCCCTGAGCAACACCCAGCAAG 1287
Db 1942 TATCTCAGAGCTTCTGGACGGCGGCGACGAGACTGTCTCCCTGGATTCCTCTGGTGG 2001
QY 1288 CCATCTCCCTCCCGAGGATGTCGGGCG-----GCCAGCTACACCTCAGCCAGCAG 1341
Db 2002 GCGCTGCGCCCTCCACACAGGCTCCGCGCGCATGGCCCTGTACCAGCTGACACGAG 2061
QY 1342 TCGAGCTGGCTTTTGGCGTGGCTCCAGCCCTGTCTTACA-----TGCAGTAC 1392
Db 2062 TCGAGCAGATCTTTTGGCGCGGATTTCCGCCACTGCCCCAACACCTCTGTCTAGGAGTC 2121
QY 1393 TGCACCAAGCTGTGGTG---CACCGGGAAGCGCAAGGACAGATGGTGTGCCAGACCGC 1449
Db 2122 TCGCGCCAGCTTGTGCTCCACTGATGGGCTGAGCCCTGTGCCACAGCAATGGC 2181
QY 1450 CACTTCCCTGGCGGATGGACCAAGCTGTGGCGAGGCAAGCTCTGCTCAAAAGGGGCC 1509
Db 2182 AGCCTGCGCTGGCTGACGGCAGCGCTGCGGGCTGGGCACTCTGCTCAGAGGCGC 2241
QY 1510 TG-----CGTGGAGAGACAACTCAACAGCACAGGCTGGATGGTCTCGTGGCCAAA 1563
Db 2242 TGTCTACCTGAGGAGAGTGGAGGCGCCAAAGCCCGTGGTAGATGGAGGCTGGGCAACG 2301
QY 1564 TGGGATCCCTATGGCCCTGCTCGCGACATGTGTGGGCGCGAGCTGAGCTGCGCAGAGG 1623
Db 2302 TGGGACCTTGGGAGATGTCTCGACCTGTGGAGGAGAGTACAGTCTACACCGT 2361
QY 1624 CAGTGCACCAACCCCTCCCAAGGGGCAAGTACTGCGAGGAGAGTCAAGGTGAAA 1683
Db 2362 GAGTGCAGAGGACCCGAGCCTCAGAATGGAGGAAGATATGCTGCTGGTGGAGAGCCAAG 2421
QY 1684 TACCGATCTGCATCTGGAGCCCTGCCCCAGCTCAGCTCCGGAAGAGCTTCGCGGAG 1743
Db 2422 TACCGATCTGCATCTGGAGCCCTGCCCCAGCTCAGCTCCGGAAGAGCTTCGCGGAG 2475
QY 1744 GAGCAGTGTAGGCTTTCAAGGCTTACACACAGCACCACCGCTCAGCTCTGCGCGTG 1803
Db 2476 CAGCAGTGTGAGAGTATATGCTTACATTTACATGACATGGAGCGGGAATCT---CCTG 2532
QY 1804 GCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGACAAAGTGCACAGCTCATCTGCCGA 1863
Db 2533 CAGTGGGTGCCAAGTATGCTGGGTGTCTCCCGGACCGCTGCAAGTCTCTGCGCA 2592
QY 1864 GCCAATGGCACTGCTACTTCTATGTGTCGACCCAGGTGGGAGGAGCAGCTGTGC 1923
Db 2593 GCGCGGGGAGGAGCGAGTCAAAAGTTCGAGGCCAAGGTGATGATGACACCTCTGT 2652
QY 1924 TCTCTGACTCCACCTCGCTGTGTCCAAAGGCAAGTGCATCAGGCTGCTGTGAGGG 1983
Db 2653 GGGCCAGAACACTGGCCATCTGTGCTGGGCGAGTGTGTCAAGGCGCGCTGTGACCAT 2712
QY 1984 AACCTGGCTCCAAAGAGATTCGAAAGTGTGGGTGTGTGGGGGAGACAAATAGAGC 2043
Db 2713 GTGTGGACTCGCTCCGAGCTGACAAATGCGGGGTGTGTGGGGGCAAGGCAACTCC 2772
QY 2044 TCGAAGAGGTGACTGAGCTTCTTCAACCAAGCCCATGATGCTGCTACATTTCTGGTGGCC 2103
Db 2773 TCGAAGAGGTGCTCGGGTCCCTCACCCCAATATGCTGCTGCTGCTGCTGCTGCTGCT 2832
QY 2104 ATCCCCCAGCGCTCAAGCATGCAATCGGCGCGGTACAAAGGCGCTGTGCTGGG 2163

Db 2833 ATCCAGCTGGTGCCACTAATATTGAGTGAAGCAGCGGAGCCACCGGGTGTGCAGAAC 2892
QY 2164 GATGACAACCTACCTGGCTCTGAAGAACAGCAGCAAGGCAAGTACCTGCTCAACGGGCAATTC 2223
Db 2893 GATGGGAACCTACCTGGCGCTGAAGACGGCTGATGGGCGAGTACCTGCTCAACGGCAACCTG 2952
QY 2224 GTGGTGTGGCGGTGGAGCGGACCTGGTGTGAAGGCGAGTCTGTGCTGCGGTACAGCGC 2283
Db 2953 GCCATCTCTGCCATAGAGCAGGACATCTTGGTGAAGGGACCATCTCTGAAGTACAGCGC 3012
QY 2284 ACGGGCACAGCGGTGGAGAGCCCTGCGAGGCTTCCGGGCCATCTCGAGCCGCTGACCGIG 2343
Db 3013 TCCATGCCACCTTGGAGCGCTGCGAGGCTTCCGGCCCTTGCAGAGCCCTCTGACAGIG 3072
QY 2344 GAGTCTCTCTCCGT---GGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTCTATCIG 2400
Db 3073 CAGCTCTGACAGTCCCTGCGAGGCTTCCGCCCAAAAGTCAATACACCTCTCTTGT 3132
QY 2401 CCCAAGA 2408
Db 3133 CCTAATGA 3140

RESULT 14

D67076

LOCUS

DEFINITION

D67076

Mouse mRNA for secretory protein containing

complete cds.

D67076

D67076.1 GI:1813339

secretory protein containing thrombospondin motifs; ADAMTS-1.

Mus musculus colon adenocarcinoma cell_line:murine colon 26 cDNA to

mRNA.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

Kuno, K., Kanada, N., Nakashima, E., Fujiki, F., Ichimura, F. and

Matsushima, K.

Molecular cloning of a gene encoding a new type of

metalloproteinase-disintegrin family protein with thrombospondin

motifs as an inflammation associated gene

J. Biol. Chem. 272 (1), 556-562 (1997)

97150761

REFERENCE

2 (bases 1 to 4180)

Kuno, K., Kanada, N., Nakashima, E., Fujiki, F., Ichimura, F. and

Matsushima, K.

Molecular cloning of a gene encoding a new type of

metalloproteinase-disintegrin family protein with thrombospondin

motifs as an inflammation associated gene

Unpublished

3 (bases 1 to 4180)

Kuno, K.

Direct Submission

Submitted (28-SEP-1995) Kouji Kuno, Cancer Research Institute,

Kanazawa University, Pharmacology; Takara-machi 13-1, Kanazawa,

Ishikawa 920, Japan (Tel:0762-62-8151(ex.5454), Fax:0762-60-7704)

Location/Qualifiers

1..4180

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="murine colon 26"

/cell_type="colon adenocarcinoma"

1..4180

/gene="ADAMTS-1"

441..3296

/gene="ADAMTS-1"

/note="putative"

/codon_start=1

/product="secretory protein containing thrombospondin

motifs"

/protein_id="BAA11088.1"

/db_xref="GI:1813340"

/translation="MGDVORAAARSGLSAHMLLLLLLLASTTLLLCARGAHRPTEEDE
ELVLSLERAPCHDSTTTLRLDADFQOOLHLKLPDSGE LAPFTLTQVRSPPGSEAO
HLDPTGDLAHCSTYGVNDGPPSAAALSCEGVGFALQGEPEFLOPAGVATERLA
PAVPEESKPSGSIIRKRVVSPRYVETMLVADQSMADFHSGSLKHLXLLTLFSVAARF
TPDQAGTRNSISLVVKKILVYIEQKGEVTSNAALTLRNFPCSWQKHNSPSDRPEHY
YKHPSTRNSISLVVKKILVYIEQKGEVTSNAALTLRNFPCSWQKHNSPSDRPEHY
DTAILFTRODLGSHSTCDTFLGMADYVTCDFPSRSCSVIEDGLQALAAFTAHLEGHVFN
MPHDDAKHCASLNGVSDSHLWASMLSSLDHSQSPSCSAYMWTFSFLDNHGECLMDK
PQNPFLKPSDLPTLYLDANROCOFTFGESKHCDDAASTCTLMCTGTSGGLLVCOIK
HPFMADGTCGGGKKVCQSKVKNKTDMDKHFATPVHGSWGPWPGWDCSRTGSGGVQYT
MRECDNPVKNGKCEGRKRVYRSCNLEDNDNNKTFREEQCAHNEFSSKSGFNE
PIVETVTPKYSRDKRLCKLCEAKGTFYFVLQPKVVDGTPCSDSTVSCVQGCVK
AGCDRTIDSKKFGCGGNGSTCKKMSGIVTSTRPGYHDIVIPAGANIEVKR
NORGSNNNSFLIAIRAADQYILNGNFTLSTLEODLYTKYVLRVSGSAAALRFSF
SPLKEPLTQVLMVGHALRPKIKFTYEMKKKTESFNAITPFESEWIEBWCSTKCS
GQWRVQCRDINGHPASCAKEVPASTRPCADLPCPHWQVGDWSPCKTCGKGYKK
RLTKVSHDGGVLSNESCDPLKKPKHYHDFCTLTQCS"

BASE COUNT 1078 a 1051 c 1146 g 905 t
ORIGIN

Query Match 22.2%; Score 634.4; DB 10; Length 4180;
Best Local Similarity 60.6%; Pred. No. 4.5e-104;
Matches 1097; Conservative 0; Mismatches 701; Indels 12; Gaps 3;

Qy 631 AAGCGTTTCGTGTATFCCCGCGGTAGTGGAGACGCTGTGTGTCGGGACGAGTCAATG 690
Db 1143 AAGCGATTTGTGTCAGCCCGGTATGTGGAACCATGCTGTAGTCAACCATCCATG 1202
Qy 691 GTCAGTTTCACGGCGGACCTTGAACATTAATCTGCTGACGCGTGTGTCGACGGCGG 750
Db 1203 GCCGACTTCCACGGCAGCGGTCTAAAGCATTAATCTTCTAAACCTGTCTCTGTTGG 1262
Qy 751 CGACTCTACCGCATCCACGATCCCTCAACCCCATCAACATCTTGTGTGTCAGAGTCTG 810
Db 1263 AGTTTTACAGCATCCACGATAGGAATTAATAGCTGTGTGTGTGTGTGTGTGTGTGT 1322
Qy 811 CTCTTTAGAGATCGTACTCGGCGCCCAAGTCAACGGCAATCGGCGCCGACGCTGCGC 870
Db 1323 GTCATATACGAGGACGAGAGGACCAAGAGTTACCTTCAACATGACGCTCTCACCTTCG 1382
Qy 871 AACTTCTGTGCTGGCAGAGAGAGCTGAACAAAGTAGTACAGACGACCCGAGTACTGG 930
Db 1383 AATTCTGACGCTGGCAGAGAACACACACAGCCCGAGTACCGGATCCAGAGCACTAT 1442
Qy 931 GACACTGCCATCTCTTACCAGGACGAGCTGTGTGGAGCCACCACTGTGACACCTG 990
Db 1443 GACACTGCAATCTGTTTACCAGACAGGATTTATGTGGCTCCACACAGCTGTGACACTC 1502
Qy 991 GCGATGGCTGATGTGGTACCATGTGTGACCCCAAGAGAGCTGCTGTGATGTAGGAC 1050
Db 1503 GGAATGGCAGATGTGGAACCGTATGTGACCCCAAGAGAGCTGCTGATGATAGAGAT 1562
Qy 1051 GATGGCTTCCATCAGCTTCAACACTGCCCACGAGCTGGGCCACGTTTCAACATGCCCC 1110
Db 1563 GATGGTTGCAAGCGCTTCCACAGAGCCCATGAATGGGCCATGTGTTTACATGCGG 1622
Qy 1111 CATGACATGTGAAGTGTGTGAGAGGTTGTGGGAAGCTTCCAGCCCAACACATGATG 1170
Db 1623 CACGATGATGCTAAGCACTGTGCCAGCTTGAATGGTGTGAGTGGCGATTCTCATCTGATG 1682
Qy 1171 TCCCGGACCTCATCAGATGACCGGTGCAACCCCTGTCAGCTGACGCTGCTGCTGCTG 1230
Db 1683 GCTTCGATGCTCTCAGCTTAGACATAGCCAGCCCTGGTCACTTTCGATGTGCTACATG 1742
Qy 1231 ATCACCAGCTTCTTCCAGCGGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290
Db 1743 GTCAGTCTTCTTCCATAGATAGGACACGGGATGTTTGTGAGACAGCCCAAGATCCA 1802
Qy 1291 ATCTCCCTGCCAGGATCTGCCGGGCGGACGCTACACCTGACCCAGCAGCTGGAGCTG 1350
Db 1803 ATCAAGCTCCCTTCTGATCTTCCCGGTACCTTGTACGATGCCACCGCCAGTGTGCT 1862
Qy 1351 GCTTTTGGCGTGGCTCCAGCCCTGCTCTTACATGATGATGATGATGATGATGATGATG 1407

```

Db      2940 AACGCCATTC 2949

RESULT 15
AF207664
LOCUS      Homo sapiens matrix metalloprotease (ADAMTSl) mRNA, complete cds.
DEFINITION
ACCESSION AF207664
VERSION    AF207664.1 GI:6685071
KEYWORDS
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 4659)
AUTHORS   Glienke,J., Schmitt,A., Pillarsky,C., Hinzmann,B., Weiss,B.,
            Rosenthal,A. and Thierauch,K.H.
TITLE     Genes differentially expressed by endothelial cells in distinct
            angiogenic states
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 4659)
AUTHORS   Glienke,J., Schmitt,A., Pillarsky,C., Hinzmann,B., Weiss,B.,
            Rosenthal,A. and Thierauch,K.H.
TITLE     Direct Submission
JOURNAL   Submitted (19-NOV-1999) Experimental Oncology, Schering AG,
            Muellerstr 178, Berlin 13342, Germany
FEATURES
            Location/Qualifiers
            1..4659
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /cell_type="endothelial cells"
            1..4659
             /gene="ADAMTSl"
            1..442
             /gene="ADAMTSl"
            443..3346
             /gene="ADAMTSl"
             /note="METH-1-like protein"
             /codon_start=1
             /product="matrix metalloprotease"
             /protein_id="AAF23772.1"
             /db_xref="GI:6685072"
             /translation="MORAVPGEGRKLGSDMGNAERAFGRSFGVPVTLTLLAALL
            AVSDALGRPSDEDELVPELERAGCHGHTTLRLHAFQCLDLLEURPOSSSELAFTLL
            QNVGRKSSETPLPDTLHCFVSGTVNGDQSSAAALSLCEGVGAFYLLGEAYFIQ
            LPAASERLALPEKDKPAPLQFHLRLNRQGVGGTGCVVDDPRPTGKAETDEDE
            GTEDEGAQSPDAPLQGVQGTGTGTSIRKRVFSHRVETMLVADQMAEPFGS
            GLKHYLLTFVSAARLYKHPSTRNSVLVVKLVIHDFQKGPVETVSNALTLRNFCS
            WQKNHTPSRDAEHYDTALFTRODLGSGTCDTPLGMADYVTCVDRSRSQSVIEDDG
            LQAATTAHELGHVFNPHDDAKOCASLGNQDSHMMASMLNLDHSQSPSCSAYM
            ITSFLDNGHGCLEMDKPNPQLQFDLPSTSDANRQCTFTGEDSKICPDAACTST
            LWCVTSGGVLCVTPFPWADGTSCGEGKWCINCKGVNKTDRKHDFPFGSGMGWF
            PWGDSRTCGGVQYTMRECDNPVKNKGKYCEGRKYRVRSCNLEDQCPDNGKTFREE
            QCAHNFEFSKASFGSGAVENIPKYAGVSPKDRCKLQCAKGIQGVFFVLPQVVDGTP
            CSPSTSVCGOQCVKAGCDRIIDSKKFKDGCYGGNGSTCKKISGVSVAKPGYHD
            IITITGATNTEVKORNSRNGSFLAIAKAADGTIILNGDYTLSTLEQDIMGV
            LRYSGSAALEPRISFSLKPELFIQVLTVGNALRPKYITFYVKKKKSENAITPFS
            AWVIEEGECSCSCELGWMRLVECDRNGQPASCEAKEVPASTRPCADHPCPQWQL
            GEWSCKTKQGVKKYKSLKCLSHDGGVLSHESCDPLKKPKHFDICTMAECS"
            3347..4659
             /gene="ADAMTSl"
BASE COUNT      1242 a 1083 c 1255 g 1079 t
ORIGIN

Query Match      22.1%; Score 629.8; DB 9; Length 4659;
Best Local Similarity 57.1%; Pred. No. 2.9e-103;
Matches 1306; Conservative 0; Mismatches 912; Indels 69; Gaps 6;

2Y      175 GCATTTCAGGAGGACTTTTACTACACTGAGCCCGGATGCTCAGTTCTTGGCTCCCGCC 234
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
dbb      680 GCCTTTGACACGACTGGATCTGGAGCTGGCGGCCGACGACGACTTTTGGCGCCGCG 739
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2Y      235 TTCTCCATGACATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGGGCTC---TTCAGAC 291

```

Db 1820 GGGAAATGTTGATGACAAAGCCTCAGAAATCCATACAGCTCCAGGCGATCCCTGGC 1879
Qy 1318 GCCAGCTACACCTGAGCAGAGTCCGAGCTGGCTTTTGGCTGGCTTCCAGCCCTGT 1377
Db 1880 ACCTCGTACAGTCCCAACCGGAGTCCGCTTTACATTTGGGAGGACCTCAACACACTGC 1939
Qy 1378 CTTTACATGACAGTAC--TGCACAAAGCTGTGTGCACCGGAAGGCCAAGGGACAGATG 1434
Db 1940 CCCGATCCAGCCAGACATGTAGCACCTTGTGTGTACCGGCACCTCTGTGGGTGTGTG 1999
Qy 1435 GTGTGCCAGACCCGCGACATTCCTCGGCGGATGGGACACAGCTGTGGCGAAGGCAAGCTC 1494
Db 2000 GTGTGTCAAAACCAACACTTCCGCTGGGCGGATGGGACAGCTGTGGAGANGGAAATGG 2059
Qy 1495 TGCTCAAGGGGCGCTGCGTGGAGAGACACAACTCAACAAG-----CACAGGTGGAT 1548
Db 2060 TGTATCAACGGCAAGTGTGTGAACAAACGGACAGAAAGCATTTTGTACGCCCTTTTCAT 2119
Qy 1549 GGTTCCTGGGCCAAATGGGATCCCTATGGCCCTGTCTCGGCGACATGTGTGGGGCGGTG 1608
Db 2120 GGAAGCTGGGGAATGTGGGGGCGCTTGGGAGACTGTTCCAGAACTGCGGTGGAGGATC 2179
Qy 1609 CAGCTGCCAGAGGAGTGCACACACCCACCCCTGCCACGGGGGCAAGTACTCGGAG 1668
Db 2180 CAGTACAGATGAGGGAATGTGACACCCAGTCCCAAGAAATGGAGGGAAGTACTGTGA 2239
Qy 1669 GGAGTGAGGCTGAAATACCGATCCTGCAATCTGGAGCCCTGCCACAGCTAGCCTCCGGA 1728
Db 2240 GGCACAGAGTGGCTACAGATCCTGTACCTTGGAGACTGTCCAGAC--NATAATGGA 2296
Qy 1729 AAGAGCTTCGGGAGGAGTGTGAGGCTTTCAAGGCTTACAAACAGCAGACCAACCCGG 1788
Db 2297 AAAACCTTTAGAGAGGAACAATGTGAAGCACACAAGAGTTTCAAAAGCTTCTTTGGG 2356
Qy 1789 CTCACCTCGCGTGGATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTGC 1848
Db 2357 AGTGGCCCTGGGTGGAAATGGATCCCAAGTACGCTGGCGCTCACCAAAAGGACAGTGC 2416
Qy 1849 AAGCTCATCTGCCGAGCAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGGTGGT 1908
Db 2417 AAGCTCATCTGCCAAGCAAAAGGCAATGGCTACTTCTCTCTTTTGCAGGCCAAAGTTGA 2476
Qy 1909 GACGGCAGCTGTGCTCTCTGACTCCACTCCGCTGTGTGTCTCCAGGCAAGTGCATCAAG 1968
Db 2477 GATGGTACTCCATGTAGCCCCAGATTCACCTCTGTCTGTGTGCAAGGACAGTGTGTAAA 2536
Qy 1969 GCTGGCTGTGATGGGAACCTCGGCTCCAAAGAAGATTCGACAAAGTGTGGGTGTGTGGG 2028
Db 2537 CTTGGTGTGATCCATCATAGACTCCAAAGAAGTGTGATAAATGTGTGTGGGG 2596
Qy 2029 GGAGACAATAAGAGCTGCAAGAAGGTGACTGGACTCTTCACCAAGCCCATGCATGGCTAC 2088
Db 2597 GGAAATGGATCTACTTGTAAAAAATATCAGGATCAGTTACTAGTGCAAAACCTGGATAT 2656
Qy 2089 AATTTGCTGGTGGCCATCCCGGAGGCGCTCAGCATCGACATCCGCCAGCGGGTTAC 2148
Db 2657 CATGATATCATCAAAATCCAACTGGAGCCCAACCAATCGAAGTGAACAGCGGAACCCAG 2716
Qy 2149 AAAGGCTGTACGGGATGACAACTACTGCTCTGAAGAACAGCCAAAGCAAGTACCTG 2208
Db 2717 AGGGATCCAGGAACAATGGCAGCTTCTTGCCATCAAGCTGCTGATGGCACATATAT 2776
Qy 2209 CTCACGGGCATTTCTGGTGTGTCGGCGGTGGAGCGGACCTGGGTGTGAAGGGAGTCTG 2268
Db 2777 CTTAATGTGACTACACTTTGTCCACCTTAGACGACAGACATATATGTACAAGGTGTGTGC 2836
Qy 2269 CTGGGTACAGCGGACAGCGGTGGAGAGCTCGAGGCTTCCCGCCCATCTCG 2328
Db 2837 TTGAGGTACAGGGCTCTCTGCGGCAATGGAAAGAAATTCGAGCTTTAGCCCTCTCAAA 2896
Qy 2329 GAGCGCTGACCGGTGGAGTCTCTCGGTGGGGAAGATGACACCGCCCGCGGTCCGCTAC 2388
Db 2897 GAGCCCTTGACCACTCCAGGTTCTTACTGTGGCAATGCCCTTTCGACCTAAAAATAATAC 2956

Qy 2389 TCCTTCT 2395
Db 2957 ACCTACT 2963

Search completed: May 15, 2003, 22:28:06
Job time : 9377 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 00:28:42 ; Search time 7221 Seconds
(without alignments)
11498.450 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtcttctgtggtcatctct.....gcgtcttgaggcgtgtctga 2853

Scoring table: OLIGO.NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 24

Total number of hits satisfying chosen parameters: 14

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_in.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_vl.*

31: em_htg_hum.*

32: em_htg_inv.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_vrt.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2802	98.2	2853	9	HSA315733	AJ315733 Homo sapi
2	2196	77.0	2930	6	AX342635	AX342635 Sequence
3	2196	77.0	2937	6	AX319860	AX319860 Sequence
4	958	33.6	157963	9	AP002986	AP002986 Homo sapi
5	958	33.6	172905	2	AC025130	AC025130 Homo sapi
6	803	28.1	170682	2	AC033429	AC033429 Homo sapi
7	724	25.4	28000	9	AP003459	AP003459 Homo sapi
8	358	12.5	172905	2	AC025130	AC025130 Homo sapi
9	288	10.1	170682	2	AC033429	AC033429 Homo sapi
10	74	2.6	182656	2	AC101990	AC101990 Mus muscu
11	61	2.1	178764	2	AC126507	AC126507 Rattus no
12	44	1.5	178764	2	AC126507	AC126507 Rattus no
13	36	1.3	182656	2	AC101990	AC101990 Mus muscu
14	35	1.2	3927	10	BC009667	BC009667 Mus muscu

ALIGNMENTS

RESULT 1
HSA315733
LOCUS
DEFINITION Homo sapiens mRNA for metalloprotease disintegrin 15 (ADAMTS15 gene).
ACCESSION AJ315733
VERSION AJ315733.1 GI:19171175
KEYWORDS ADAMTS15 gene; disintegrin; metalloprotease; thrombospondin.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cal, S., Obaya, A.J., Llamazares, M., Garabaya, C., Quesada, V. and Lopez-Otin, C.
TITLE Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains
JOURNAL Gene 283 (1-2), 49-62 (2002)
MEDLINE 21856482
PUBMED 11867212
REFERENCE 2 (bases 1 to 2853)
AUTHORS Cal, S.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2001) Cal S., Biochemistry and Molecular Biology, University of Oviedo, Campus del Cristo, Asturias. 33006, SPAIN
FEATURES
source
1..2853
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2853
/gene="ADAMTS15"
1..2853
/gene="ADAMTS15"
/codon_start=1
/product="metalloprotease disintegrin 15 with thrombospondin domains"
/protein_id="CAC86014.1"
/db_xref="GI:19171175"
/translation="MLLGLTLAFAGRTAGGSEPEVVPRLIOPDINGRRYYWRG
PEDSGDGLIFQITAFQEDFYHLTPDAQFLAPAFSTELGVPLAGLTGSGSDLPKRCF
YSGDYNAPDSEFAVSLCGGLRGAFYGAETVISPENAPAAQRNSQAHLLQRR
GVPGGSDPTSCRGVAGSNPAILRLADLPYKPRRAGGSESRSSRRKRRFVSIPR
YVETLVADSESNVHFQADLEHLTLTALATARYRHPESILNPINIVVKVLLLRDR
SGPKVTGNAALTFRNFKCAWKLNKVSQDKHPYWDTAIFLTRQDLCGATCTDGLGMD
VGTMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVEVFCGLRANHMSP

TLIQIDRANPWSACSAIIITDFLDSHGDCLLDQPSKPISLPEDIPLGASITUSQOCEL
AFVGSKPCPYMOYCKLWCTKAKGQMVQCTRHFPWADGTSCGSKLCLKAGACVERH
NLNKHVDSGSAKWDYGPCSRTCTGGVQLARRQTNPTPANGKYCEGVRYKRSN
LEPCPSASGKSPREQCEAFNGYHNHNRKTLAVAWVPKYSVSPRDKCKLICRANG
TGYYVLAPKVDGTLTCSPDSTSVQVQKCIKAGDGNLSKRRDKGCVCGDNKSC
KVTGLETPEMGYNFVWALPAGASSIDIRQRYKGLIGDDNYLALKNSOGKXLLNGH
FVVSVERDLVWGSLLRISGSGTAVESLQASRPILLEPTEVLSVGMTPPRVYSF
YLPKEPREDKSHPKDPRGFSVLHNSVLSNOVEQDDPRPARVWAGSWGPCSASCG
SGLQKRAVRCRSGQRTVPACDAHRPVETQACGEPCTWELSAWSPCSKSGRGFQ
RRSLKCVGHGRLLRDQNLHRKPOELDFCVLRPC"
BASE COUNT 522 a 921 c 903 g 507 t
ORIGIN

Query Match 98.2%; Score 2802; DB 9; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGCATCTTAACCTGGCTTTCGCGGGCGAACCCTGGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGCATCTTAACCTGGCTTTCGCGGGCGAACCCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGAGTAGTCTTCCATCCGACTGGACCCGGACATTAACGSCCCGCCCTAC 120
DB 61 CCAGAGCGGAGGAGTAGTCTTCCATCCGACTGGACCCGGACATTAACGSCCCGCCCTAC 120
QY 121 TACTGGCGGGGTCCCGAGGACTCCGGGGATCCGGGACTCAITTTTCAGATCAGACATTT 180
DB 121 TACTGGCGGGGTCCCGAGGACTCCGGGGATCCGGGACTCAITTTTCAGATCAGACATTT 180
QY 181 CAGGAGACTTTTACCTACACCTTGACCCCGGATGCTCAGTCTTTGGCTCCCGCTTCTCC 240
DB 181 CAGGAGACTTTTACCTACACCTTGACCCCGGATGCTCAGTCTTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATGTGGCGTCCCGCTCCAGGGCTCACCGGGGCTCTTACAGACTTCGACGC 300
DB 241 ACTGAGCATGTGGCGTCCCGCTCCAGGGCTCACCGGGGCTCTTACAGACTTCGACGC 300
QY 301 TGCCTTCTATCTGGGACGTAAGCCCGAGCGGACTCGTTCGCTGCTGTGAGCCCTGTGC 360
DB 301 TGCCTTCTATCTGGGACGTAAGCCCGAGCGGACTCGTTCGCTGCTGTGAGCCCTGTGC 360
QY 361 GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGGCCGAGTATGTCAATTAGCCCGCTGCC 420
DB 361 GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGGCCGAGTATGTCAATTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGCGGCGAGCAAGCCAGCGGCGCACACCTTCTCCAGCGCGGG 480
DB 421 AATGCTAGCGCGCGGGCGGCGAGCAAGCCAGCGGCGCACACCTTCTCCAGCGCGGG 480
QY 481 GGTCTTCCGGCGGGCTTCCGAGAGCCCACTCTCGCTCGGGGTGGCTTCGGGCTGG 540
DB 481 GGTCTTCCGGCGGGCTTCCGAGAGCCCACTCTCGCTCGGGGTGGCTTCGGGCTGG 540
QY 541 AACCCCGCCATCTTACGGGCGCTTGACCCCTTACAAGCCGGCGGGCGGCGGCTTCGGGGAG 600
DB 541 AACCCCGCCATCTTACGGGCGCTTGACCCCTTACAAGCCGGCGGGCGGCGGCTTCGGGGAG 600
QY 601 AGTGTAGCGCGGCGAGGCTTGGCGGCGCAAGGGTTTCGTCTATCCCGCGGTACGTG 660
DB 601 AGTGTAGCGCGGCGAGGCTTGGCGGCGCAAGGGTTTCGTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGTGTGCGGGAGAGTCAATGTTCAAGTTCCACGGCGGAGACCTTGGAAACAT 720
DB 661 GAGACGCTGTGTGCGGGAGAGTCAATGTTCAAGTTCCACGGCGGAGACCTTGGAAACAT 720
QY 721 TATCTGTGACGCTGTGGCAAGGGCGGCGACTCTACCGCCATCCAGACTCCCTCAAC 780
DB 721 TATCTGTGACGCTGTGGCAAGGGCGGCGACTCTACCGCCATCCAGACTCCCTCAAC 780
QY 781 CCCATCAACATCGTGTGGTCAAGGTCGCTCTTAGAGATCGTGNATCCGGGCGCAAG 840
DB 781 CCCATCAACATCGTGTGGTCAAGGTCGCTCTTAGAGATCGTGNATCCGGGCGCAAG 840

QY 841 GTCACCGGCAATGCGGCGCTGACGCTGCGCAACTTCTCTGTGCTGGCAGAAAGCTGAAC 900
DB 841 GTCACCGGCAATGCGGCGCTGACGCTGCGCAACTTCTCTGTGCTGGCAGAAAGCTGAAC 900
QY 901 AAAGTAGTGACAACACCCCGAGTACTGGGACACTGCATCTCTTCCACAGGCAAGAC 960
DB 901 AAAGTAGTGACAACACCCCGAGTACTGGGACACTGCATCTCTTCCACAGGCAAGAC 960
QY 961 CTGTGTGGAGCCACCACTGTGTGACACCCCTGGCATGGCTGTGTGGTACCATGTGTGAC 1020
DB 961 CTGTGTGGAGCCACCACTGTGTGACACCCCTGGCATGGCTGTGTGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGCTGTCTGTCTATTGAGACAGATGGGCTTCCATCAGCCTTCCACACTGCC 1080
DB 1021 CCCAAGAGAAGCTGTCTGTCTATTGAGACAGATGGGCTTCCATCAGCCTTCCACACTGCC 1080
QY 1081 CACGAGCTGGGCGACGTGTTCAACATGCCCATGACATGTGAAAGTCTGTGAGGAGTG 1140
DB 1081 CACGAGCTGGGCGACGTGTTCAACATGCCCATGACATGTGAAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGACCCCTCATCCAGATCGAOCGTGCC 1200
DB 1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGACCCCTCATCCAGATCGAOCGTGCC 1200
QY 1201 AACCCCTGGTACAGCTGCAGTCTGCCATCATCAGCAGTCTCCTGGACAGCGGCAAGGT 1260
DB 1201 AACCCCTGGTACAGCTGCAGTCTGCCATCATCAGCAGTCTCCTGGACAGCGGCAAGGT 1260
QY 1261 GACTGCTCTCTGGACAAACCCAGCAAGCCCATCTCCCTGCCCCGAGGATCTGCCGGGTGCC 1320
DB 1261 GACTGCTCTCTGGACAAACCCAGCAAGCCCATCTCCCTGCCCCGAGGATCTGCCGGGTGCC 1320
QY 1321 AGCTACACCTTGAGCGACGAGTGGCTGCTTTTGGCGTGGGCTTCCAAAGCCTGTGCT 1380
DB 1321 AGCTACACCTTGAGCGACGAGTGGCTGCTTTTGGCGTGGGCTTCCAAAGCCTGTGCT 1380
QY 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGCCAAAGGACACATGTTGTGC 1440
DB 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGCCAAAGGACACATGTTGTGC 1440
QY 1441 CAGACCCGCCACTTCCCTTGGGCGCATGACACAGCTGTGGCGAGGCAAGCTCTGCGTC 1500
DB 1441 CAGACCCGCCACTTCCCTTGGGCGCATGACACAGCTGTGGCGAGGCAAGCTCTGCGTC 1500
QY 1501 AAAGGGGCTGGTGGAGAGACAACTCAACAGCAGAGGCTGGATGGTTCCTGGGCG 1560
DB 1501 AAAGGGGCTGGTGGAGAGACAACTCAACAGCAGAGGCTGGATGGTTCCTGGGCG 1560
QY 1561 AAATGGGATCCCTATGGCCCTGCTCGCGCACATGTGTGGGGCGTGCAGCTGGCGAG 1620
DB 1561 AAATGGGATCCCTATGGCCCTGCTCGCGCACATGTGTGGGGCGTGCAGCTGGCGAG 1620
QY 1621 AGGAGTGCACCAACCCCAACCCCTGCCAACGGGGCAAGTACTGCGAGGAGTGAAGTG 1680
DB 1621 AGGAGTGCACCAACCCCAACCCCTGCCAACGGGGCAAGTACTGCGAGGAGTGAAGTG 1680
QY 1681 AAATACCGATCTGCAATCTGAGCCCTGCCAGCTCAGCTCGGAAAGAGCTTCGG 1740
DB 1681 AAATACCGATCTGCAATCTGAGCCCTGCCAGCTCAGCTCGGAAAGAGCTTCGG 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAGCAGCCGCTCAGCTTCGCC 1800
DB 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAGCAGCCGCTCAGCTTCGCC 1800
QY 1801 GTGGCATGGTGGCCCAAGTACTTCCGGCGTGTCTCCCGGGACAAAGTGAAGCTCATGTC 1860
DB 1801 GTGGCATGGTGGCCCAAGTACTTCCGGCGTGTCTCCCGGGACAAAGTGAAGCTCATGTC 1860
QY 1861 CGAGCAATGGCAGTGGCTTCTTATGTGTGACCCCAAGGTTGGAGCGGACGCTG 1920
DB 1861 CGAGCAATGGCAGTGGCTTCTTATGTGTGACCCCAAGGTTGGAGCGGACGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCCGCTGTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGAT 1980

Db 1921 TGCTCTCTCTGACTCCACCTCCGCTGTGTCCAAAGGCAAGTCAATCAAGGCTGGCTGTGAT 1980
QY 1981 GGAACCTGGCTCCCAAGAGAGATTCCACAAGTGTGGGGTGTGTGGGGGAGACAATAAG 2040
Db 1981 GGAACCTGGCTCCCAAGAGAGATTCCACAAGTGTGGGGTGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAAGTGTGACTGGACTCTTCCACCAAGCCCATGCTACATGCTCAATTTCTGGTGTG 2100
Db 2041 AGCTGCAAGAAGTGTGACTGGACTCTTCCACCAAGCCCATGCTACATGCTCAATTTCTGGTGTG 2100
QY 2101 GCCATCCCGCAGGCGCTCAAGCATGACATCCGCGCAGCGGTTTCAAAGGGCTGATC 2160
Db 2101 GCCATCCCGCAGGCGCTCAAGCATGACATCCGCGCAGCGGTTTCAAAGGGCTGATC 2160
QY 2161 GGGGATGACAACTACTCGGCTCTGAAGAAGCAAGCAAGCAAGTACCTGCTCAAGCGGCAT 2220
Db 2161 GGGGATGACAACTACTCGGCTCTGAAGAAGCAAGCAAGCAAGTACCTGCTCAAGCGGCAT 2220
QY 2221 TTGCTGTGTGTCGGCGGTGAGCGGACCTGTGTGTGAAGGCACTCTGCTCGGTACAGC 2280
Db 2221 TTGCTGTGTGTCGGCGGTGAGCGGACCTGTGTGTGAAGGCACTCTGCTCGGTACAGC 2280
QY 2281 GGCAGGCGCAGCGGTGAGAGCCTTGAGGCTTCCGCGGCCATCTCTGGAGCGGCTGACC 2340
Db 2281 GGCAGGCGCAGCGGTGAGAGCCTTGAGGCTTCCGCGGCCATCTCTGGAGCGGCTGACC 2340
QY 2341 GTGAGGTCTCTCTCGGTGGGGAAGATGACACCGCGCGGTCTGCTACTCTCTCTATCTG 2400
Db 2341 GTGAGGTCTCTCTCGGTGGGGAAGATGACACCGCGCGGTCTGCTACTCTCTCTATCTG 2400
QY 2401 CCCAAGAGCTCGGAGGAGCAAGTCTCTATCCCAAGGACCTCCGCGGACCCCTCTGTC 2460
Db 2401 CCCAAGAGCTCGGAGGAGCAAGTCTCTATCCCAAGGACCCCGCGGACCCCTCTGTC 2460
QY 2461 TTGCACAACAGCGCTCAGCTCTCCCAAGAGTGGAGCAGCGGAGCAGCGCCCT 2520
Db 2461 TTGCACAACAGCGCTCAGCTCTCCCAAGAGTGGAGCAGCGGAGCAGCGCCCT 2520
QY 2521 GCACGCTGGGTGGTGGCAGCTGGGGCGCTGTCTCCGCGAGCTGGCGAGTGGCGTGCAG 2580
Db 2521 GCACGCTGGGTGGTGGCAGCTGGGGCGCTGTCTCCGCGAGCTGGCGAGTGGCGTGCAG 2580
QY 2581 AAGCGCGGTGGACTGTGGGGTTCGCGCGGAGCGCAGGTCGCTGTGTATGCA 2640
Db 2581 AAGCGCGGTGGACTGTGGGGTTCGCGCGGAGCGCAGGTCGCTGTGTATGCA 2640
QY 2641 GCCATCGGCGCTGGAGACACAGCTGGGGGAGCGCTGGGGAGCGCTGGGGAGCTCAGC 2700
Db 2641 GCCATCGGCGCTGGAGACACAGCTGGGGGAGCGCTGGGGAGCGCTGGGGAGCTCAGC 2700
QY 2701 GCCTGTACCTGTCTCAAGAGCTGGCGCGGGGATTTTCAGAGCGGCTCACTCAAGTGT 2760
Db 2701 GCCTGTACCTGTCTCAAGAGCTGGCGCGGGGATTTTCAGAGCGGCTCACTCAAGTGT 2760
QY 2761 GTGGGCGCAGCGGCTGTGGCGCGGAGCAGTGGAGCTGCACTGGAGCGGCGGAG 2820
Db 2761 GTGGGCGCAGCGGCTGTGGCGCGGAGCAGTGGAGCTGCACTGGAGCGGCGGAG 2820
QY 2821 GAGCTGACTTCTCGCTCTGAGCGGCTGTGA 2853
Db 2821 GAGCTGACTTCTCGCTCTGAGCGGCTGTGA 2853

RESULT 2
AX342635
LOCUS
DEFINITION
Sequence 32 from Patent WO0198468.
ACCESSION
AX342635
VERSION
AX342635.1
KEYWORDS
GI:18152032
SOURCE
human.
ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Yue, H., Elliott, V. S., Gandhi, A. R., Lal, P., Au-Young, J.,
Tribouley, C. M., Deleage, A. M., Baughn, M. R., Nguyen, D. B., Lee, E. A.,
Hafalia, A., Khan, F. A., Wallis, N. K., Yao, M. G., Lu, D. A., Patterson, C.,
Tang, Y. T., Walsh, R. T., Azimzai, I., Ramkumar, J., Xu, Y., and Reddy, R.
Patent: WO 0198468-A 32 27-DEC-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1..2930
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7473089CBI"
BASE COUNT 529 a 958 c 931 g 512 t
ORIGIN
Query Match 77.0%; Score 2196; DB 6; Length 2930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 ATGCTTCTCTGGGCACTCTAAACCTTGGCTTTGCGCGGCGGAACCGCTCGAGTCTGTGAG 60
Db 75 ATGCTTCTCTGGGCACTCTAAACCTTGGCTTTGCGCGGCGGAACCGCTCGAGTCTGTGAG 134
QY 61 CCAGAGCGGAGGTAGTCTGCTCCATCCGACTGGACCGGAGCATTAACGGCGCGGCTAC 120
Db 135 CCAGAGCGGAGGTAGTCTGCTCCATCCGACTGGACCGGAGCATTAACGGCGCGGCTAC 194
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGACTCATTTTTCAGATCACAGTATT 180
Db 195 TACTGGCGGGTCCCGAGGACTCCGGGGATCAGGACTCATTTTTCAGATCACAGTATT 254
QY 181 CAGGAGGACTTTTACTACCTGACGCGGATGCTAGTCTTGGCTCCCGCTTTC 240
Db 255 CAGGAGGACTTTTACTACCTGACGCGGATGCTAGTCTTGGCTCCCGCTTTC 314
QY 241 ACTGAGCATCTGGGGTCCCGCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTGACGC 300
Db 315 ACTGAGCATCTGGGGTCCCGCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTGACGC 374
QY 301 TCGCTTATTTCTGGGAGCTGAACCGGAGCTGCTGCTGCTGTGTAGGCTGTGC 360
Db 375 TCGCTTATTTCTGGGAGCTGAACCGGAGCTGCTGCTGCTGTGTAGGCTGTGC 434
QY 361 GGGGGCTCCCGGAGCCTTTGGCTACGAGCGCGGAGTATGCTATGAGCGGCTGCC 420
Db 435 GGGGGCTCCCGGAGCCTTTGGCTACGAGCGCGGAGTATGCTATGAGCGGCTGCC 494
QY 421 AATGCTAGCGCGCGGCGCAGCGCAACAGCGGAGCTTCTTCAGGCGCGG 480
Db 495 AATGCTAGCGCGCGGCGCAGCGCAACAGCGGAGCTTCTTCAGGCGCGG 554
QY 481 GGTGTTCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCGCTCGGCTGG 540
Db 555 GGTGTTCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCGCTCGGCTGG 614
QY 541 AACCCCGCATCTTACGCGGCTTACAGCGCGGCGGCGGCGGCTTCGGGGAG 600
Db 615 AACCCCGCATCTTACGCGGCTTACAGCGCGGCGGCGGCGGCTTCGGGGAG 674
QY 601 AGTGTAGCGCGCGGAGGTCTGGGCGCGCAAGGCTTTCGCTGTATCCCGCGGTAGCTG 660
Db 675 AGTGTAGCGCGCGGAGGTCTGGGCGCGCAAGGCTTTCGCTGTATCCCGCGGTAGCTG 734
QY 661 GAGAGCTGTGTGCGGAGAGTCAATGTGCAAGTTCACGCGCGGAGCTCGAAT 720
Db 735 GAGAGCTGTGTGCGGAGAGTCAATGTGCAAGTTCACGCGCGGAGCTCGAAT 794
QY 721 TATCTGTGACGCTGTGGCAACGCGGCGGAGCTTACCGCCATCCAGCATCTTCAAC 780
Db 795 TATCTGTGACGCTGTGGCAACGCGGCGGAGCTTACCGCCATCCAGCATCTTCAAC 854

```
QY 781 CCCATCAACATCGTTGGTCAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
Db 855 CCCATCAACATCGTTGGTCAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 914
QY 841 GTCACCGCAATCGGGCCCTGACCTGGCAACTTCTGCTGCCGCAAGAGACGTCAAC 900
Db 915 GTCACCGCAATCGGGCCCTGACCTGGCAACTTCTGCTGCCGCAAGAGACGTCAAC 974
QY 901 AAAGTGAGTGACAAGCACCCTGAGTACTGGGACACTGCCATCTCTTCCACGACGAC 960
Db 975 AAAGTGAGTGACAAGCACCCTGAGTACTGGGACACTGCCATCTCTTCCACGACGAC 1034
QY 961 CTGTGTGAGGACCAACCTGTGACACCTGGGCATGCTGATGTGGTATCATGTGTGAC 1020
Db 1035 CTGTGTGAGGACCAACCTGTGACACCTGGGCATGCTGATGTGGTATCATGTGTGAC 1094
QY 1021 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGTTCATCAGGCTTCAACCACTGCC 1080
Db 1095 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGTTCATCAGGCTTCAACCACTGCC 1154
QY 1081 CACGAGCTGGGCCAGTGTTCACATGCCCATGACCAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1155 CACGAGCTGGGCCAGTGTTCACATGCCCATGACCAATGTGAAGTCTGTGAGGAGGTG 1214
QY 1141 TTGTGGAAAGCTCCGAGCAACCATGATGTCGCCGACCTCATCCAGATCGACCGTGC 1200
Db 1215 TTGTGGAAAGCTCCGAGCAACCATGATGTCGCCGACCTCATCCAGATCGACCGTGC 1274
QY 1201 AACCCCTGCTGAGCTGCAGTGTGCTCATCATCACCGACTTCTTGGACAGCGGACAGGT 1260
Db 1275 AACCCCTGCTGAGCTGCAGTGTGCTCATCATCACCGACTTCTTGGACAGCGGACAGGT 1334
QY 1261 GACTGCCCTCTGGACCAACCCAGCAAGCCATCTCCCTGCCGAGGATCTGCCGGGGGCC 1320
Db 1335 GACTGCCCTCTGGACCAACCCAGCAAGCCATCTCCCTGCCGAGGATCTGCCGGGGGCC 1394
QY 1321 AGCTACACCTGAGCCAGCTGCGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCT 1380
Db 1395 AGCTACACCTGAGCCAGCTGCGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCT 1454
QY 1381 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1440
Db 1455 TACATGCACTACTGCACCAAGCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1514
QY 1441 CAGACCCGCCACTTCCCTGGGCGATGGCACAGCTGTGCGGAGGCAAGCTCTGCTC 1500
Db 1515 CAGACCCGCCACTTCCCTGGGCGATGGCACAGCTGTGCGGAGGCAAGCTCTGCTC 1574
QY 1501 AAAGGGGCTGCTGGAGAGACACAACCTCAACAAGCACAGGCTGGATGTTCTTGGGCC 1560
Db 1575 AAAGGGGCTGCTGGAGAGACACAACCTCAACAAGCACAGGCTGGATGTTCTTGGGCC 1634
QY 1561 AAATGGGATCCCTATGCGCCCTGTGCGGCACATGTGTGGGGGCGTGCAGCTGGCCAGG 1620
Db 1635 AAATGGGATCCCTATGCGCCCTGTGCGGCACATGTGTGGGGGCGTGCAGCTGGCCAGG 1694
QY 1621 AGGCAGTGCACCAACCCCTCCCAACCGGGGCAAGTACTCCGAGGAGTGAAGGTG 1680
Db 1695 AGGCAGTGCACCAACCCCTCCCAACCGGGGCAAGTACTCCGAGGAGTGAAGGTG 1754
QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGGCTTCGGAAAGAGCTTCCGG 1740
Db 1755 AAATACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGGCTTCGGAAAGAGCTTCCGG 1814
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAAACCACAGCAACACCGGCTCACTCTGCC 1800
Db 1815 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAAACCACAGCAACACCGGCTCACTCTGCC 1874
QY 1801 GTGGCATGGTCCCAAGTACTCCGGGTGTCTCCCGGGACAGTGCACAGCTCATCTGC 1860
Db 1875 GTGGCATGGTCCCAAGTACTCCGGGTGTCTCCCGGGACAGTGCACAGCTCATCTGC 1934
QY 1861 CGAGCAATGGCACTGGCTACTTCTATGTGTGGCACCCCAA ---GGTGGTGGACGGCAGC 1917
```

```
Db 1935 CGAGCAATGGCACTGGCTACTTCTATGTGTGGCACCCCAAGGTGGTGGACGCTACG 1994
QY 1918 CTGTGCTCTCCTGACTCCACCTCCGCTGTGTCCAAAGCAAGTGCATCAAGGCTGCTGT 1977
Db 1995 CTGTGCTCTCCTGACTCCACCTCCGCTGTGTCCAAAGCAAGTGCATCAAGGCTGCTGT 2054
QY 1978 GATGGGAACCTGGGCTCCAAAGAGATTCGACAAGTGTGGGGTGTGTGGGGGAGACAAT 2037
Db 2055 GATGGGAACCTGGGCTCCAAAGAGATTCGACAAGTGTGGGGTGTGTGGGGGAGACAAT 2114
QY 2038 AAGAGTGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2097
Db 2115 AAGAGTGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2174
QY 2098 GTGGCCATCCCGCAGCGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGGGCTG 2157
Db 2175 GTGGCCATCCCGCAGCGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGGGCTG 2234
QY 2158 ATCGGGGATGACAATACCTGCTCTGAGAACACAGCAAGCAAGTACCTGCTCAACGGG 2217
Db 2235 ATCGGGGATGACAATACCTGCTCTGAGAACACAGCAAGCAAGTACCTGCTCAACGGG 2294
QY 2218 CATTTCTGCTGTGCGCGGTGGAGCGGACCTGTGTGAGGGGAGTGTGTGCGGTAC 2277
Db 2295 CATTTCTGCTGTGCGCGGTGGAGCGGACCTGTGTGAGGGGAGTGTGTGCGGTAC 2354
QY 2278 AGCGGCACGGGCACAGCGGTGAGAGCTGTGAGGCTTCCCGGCCCATCTTGAGGCGGCTG 2337
Db 2355 AGCGGCACGGGCACAGCGGTGAGAGCTGTGAGGCTTCCCGGCCCATCTTGAGGCGGCTG 2414
QY 2338 ACCTGGAGGCTCTCTCCGTTGGGAGAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2397
Db 2415 ACCTGGAGGCTCTCTCCGTTGGGAGAGATGACACCGCCCGGGTCCGCTACTCTTCTAT 2474
QY 2398 CTGCCCCAAAGAGCTCGGGAGGACAGTCTCTCATCCC 2436
Db 2475 CTGCCCCAAAGAGCTCGGGAGGACAGTCTCTCATCCC 2513

RESULT 3
AX319860
LOCUS AX319860 2937 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 24 from Patent WO0183782.
ACCESSION AX319860
VERSION AX319860.1 GI:17901450
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and
Payne,V.
TITLE Novel proteases
JOURNAL Patent: WO 0183782-A 24 08-NOV-2001;
Sugen, Inc. (US)
FEATURES
source
1. .2937
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 531 a 958 c 936 g 512 t
ORIGIN

Query Match 77.0%; Score 2196; DB 6; Length 2937;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 ATGCTTCTGCTGGGCACTCAACCTTGGCTTTTCGGCGGGCAACCGCTGGAGGCTCTGAG 60
Db 76 ATGCTTCTGCTGGGCACTCAACCTTGGCTTTTCGGCGGGCAACCGCTGGAGGCTCTGAG 135
QY 61 CCAGAGCGGAGGTAGTCTGTTCCCATCCGATCCGAGTGGACCCCGGACATTAACGGCCCGCTAC 120
```

Db	136	CCAGACGGGAGGTAGTCTGCTTCCCATCGAGCTGGACCCGGACATTAACGGCCGCGGTAC	195
QY	121	TACTGCGGGTCCCGAGGACTCCGGGATCAGGAGACATATTTTCAGATCACAGCATTT	180
Db	196	TACTGCGGGGTCCCGAGGACTCCGGGATCAGGAGCATATTTTCAGATCACAGCATTT	255
QY	181	CAGGAGGACTTTTACTATACACTGACGCCGGATGCTCAGTTCCTTGGCTCCGCTTCTCC	240
Db	256	CAGGAGGACTTTTACTATACACTGACGCCGGATGCTCAGTTCCTTGGCTCCGCTTCTCC	315
QY	241	ACTGAGCATCTGGCGTCCCTCCAGGGCTCACCGGGGCTCTCAGACCTTCGACACG	300
Db	316	ACTGAGCATCTGGCGTCCCTCCAGGGCTCACCGGGGCTCTCAGACCTTCGACACG	375
QY	301	TGCTTCTATTCTGGGACGTGAACCCGAGCCGACCTCGTTCGCTGCTGTGAGCTGTGC	360
Db	376	TGCTTCTATTCTGGGACGTGAACCCGAGCCGACCTCGTTCGCTGCTGTGAGCTGTGC	435
QY	361	GGGGGCTCCGGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCAATTAGCCGCTGCC	420
Db	436	GGGGGCTCCGGGAGCCTTTGGCTACCGAGGCGCGAGTATGTCAATTAGCCGCTGCC	495
QY	421	AATGCTAGGCGCGCGCGCAGCGCAACGCCAGGCGCACACCTTCTCCAGCGCGG	480
Db	496	AATGCTAGGCGCGCGCGCAGCGCAACGCCAGGCGCACACCTTCTCCAGCGCGG	555
QY	481	GGTGTTCGGGGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGTCTCGGGCTGG	540
Db	556	GGTGTTCGGGGGGGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGGTGG	615
QY	541	AACCCCGCATCTACGGGCCCTTGGACCCCTTACAAGCCGCGCGGGGCTTCGGGGAG	600
Db	616	AACCCCGCATCTACGGGCCCTTGGACCCCTTACAAGCCGCGCGGGGCTTCGGGGAG	675
QY	601	AGTCGTAGCGGGCGCAGGTCTGGGGCGGCCAAGCGTTTCGTGTCTATCCCGGCTAGTG	660
Db	676	AGTCGTAGCGGGCGCAGGTCTGGGGCGGCCAAGCGTTTCGTGTCTATCCCGGCTAGTG	735
QY	661	GAGACGCTGTGTCTCGGACGAGTCAATGTGTCAAGTTCACGGCGGAGCTCGGAACAT	720
Db	736	GAGACGCTGTGTCTCGGACGAGTCAATGTGTCAAGTTCACGGCGGAGCTCGGAACAT	795
QY	721	TATCTGCTGACGCTGTGGCAACGGCGGCGGACCTCTACCGCCATCCAGCATCCTCAAC	780
Db	796	TATCTGCTGACGCTGTGGCAACGGCGGCGGACCTCTACCGCCATCCAGCATCCTCAAC	855
QY	781	CCCATCAACATCTTGTGTCTGATGAGTGTCTTCTAGAGATCGTAGTCTCGGGCCCAAG	840
Db	856	CCCATCAACATCTTGTGTCTGATGAGTGTCTTCTAGAGATCGTAGTCTCGGGCCCAAG	915
QY	841	GTACCCGGCAATCGCCCTGACGCTGCGCAACTCTGTTCCTGGCAGAAAGCTGAAC	900
Db	916	GTACCCGGCAATCGCCCTGACGCTGCGCAACTCTGTTCCTGGCAGAAAGCTGAAC	975
QY	901	AAAGTGAAGTCAACAGCACCCGAGTACTGGGACACTGCCATCCTTCTCACTAGGCAGGAC	960
Db	976	AAAGTGAAGTCAACAGCACCCGAGTACTGGGACACTGCCATCCTTCTCACTAGGCAGGAC	1035
QY	961	CTGTGTGGAGCCACCACTGTACACCTTGGGCATGGGTGATGTGGGTACCATGTGTAC	1020
Db	1036	CTGTGTGGAGCCACCACTGTACACCTTGGGCATGGGTGATGTGGGTACCATGTGTAC	1095
QY	1021	CCCAAGAGAAGTGTCTGTCAATTGAGGACGATGGGCTTCCATCAGCCTTACACACATGCC	1080
Db	1096	CCCAAGAGAAGTGTCTGTCAATTGAGGACGATGGGCTTCCATCAGCCTTACACACATGCC	1155
QY	1081	CACGAGCTGGGCCACAGCTGTTCACATGCCCCATGACATGTGAAGTCTGTGAGGAGGTG	1140
Db	1156	CACGAGCTGGGCCACAGCTGTTCACATGCCCCATGACATGTGAAGTCTGTGAGGAGGTG	1215
QY	1141	TTTGGGAAGTCCGAGGCCAACCATGATGTCCCGGACCCCTCATCCAGATGACCGGTGCC	1200

```
QY 2278 AGCGCACGGGCACAGCGGTGGAGAGCTGCAGGCTTCCGGGCCCATCTCGAGCGCGTG 2337
|||||
Db 2356 AGCGCACGGGCACAGCGGTGGAGAGCTGCAGGCTTCCGGGCCCATCTCGAGCGCGTG 2415
|||||
QY 2338 ACCGTGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTAT 2397
|||||
Db 2416 ACCGTGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTAT 2475
|||||
QY 2398 CTGCCAAAGAGCTCGGGAGGACAAGTCTCTCATCCC 2436
|||||
Db 2476 CTGCCAAAGAGCTCGGGAGGACAAGTCTCTCATCCC 2514
|||||

RESULT 4
AP002986 157963 bp DNA linear PRI 08-MAR-2002
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-121M22,
DEFINITION complete sequence.
ACCESSION AP002986
VERSION AP002986.2 GI:19263031
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-121M22.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2000)
2 (bases 1 to 157963)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (01-DEC-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Mar 7, 2002 this sequence version replaced gi:11559301.
FEATURES
Location/Qualifiers
1..157963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-121M22"

BASE COUNT 39742 a 36722 c 37247 g 44252 t
ORIGIN

Query Match 33.6%; Score 958; DB 9; Length 157963;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
|||||
Db 134115 ATGCTTCTGCTGGGATCCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 134174
|||||
QY 61 CCAGAGCGGAGGTAGTCTCCATCCGACTGGACCGGACATTAACGGCCGCGGTAC 120
|||||
Db 134175 CCAGAGCGGAGGTAGTCTCCATCCGACTGGACCGGACATTAACGGCCGCGGTAC 134234
|||||
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT 180
|||||
Db 134235 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT 134294
|||||
QY 181 CAGGAGGACTTTACCTACACTGACCGCGATGCTAGTCTTGGCTCCCGCTTCTCC 240
|||||
Db 134295 CAGGAGGACTTTACCTACACTGACCGCGATGCTAGTCTTGGCTCCCGCTTCTCC 134354
|||||
QY 241 ACTGAGCATCTGGGGTCCCTCCAGGGCTCCACGGGGCTCTTCAGACCTCGCAGCG 300
|||||
```

```
Db 134355 ACTGAGCATCTGGGGCTCCCCCTCCAGGGGCTACCGGGGGCTCTTCAGACCTCGCAGCG 134414
|||||
QY 301 TGCCTTCTATTTGGGGACGTGAAGCCCGAGCCGAGCTCGTTTCGCTGCTGTGAGCTCTGC 360
|||||
Db 134415 TGCCTTCTATTTGGGGACGTGAAGCCCGAGCCGAGCTCGTTTCGCTGCTGTGAGCTCTGC 134474
|||||
QY 361 GGGGGGCTCCCGGGAGCCTTTGGGTACCGAGGCGCGCGAGTATGTCATTAGCCCGCTGCC 420
|||||
Db 134475 GGGGGGCTCCCGGGAGCCTTTGGGTACCGAGGCGCGCGAGTATGTCATTAGCCCGCTGCC 134534
|||||
QY 421 AATGCTAGCGCGCGCGCGCGAGCAACAGCCAGGCGCACACCTTCTCCAGCGCGG 480
|||||
Db 134535 AATGCTAGCGCGCGCGCGCGAGCAACAGCCAGGCGCACACCTTCTCCAGCGCGG 134594
|||||
QY 481 GGTGTTCCGGCGGGCTTCCGGGAGACCCACCTCTCTCGCTCGGGGTGGCTTCGGGCTGG 540
|||||
Db 134595 GGTGTTCCGGCGGGCTTCCGGGAGACCCACCTCTCTCGCTCGGGGTGGCTTCGGGCTGG 134654
|||||
QY 541 AACCCCGGCATCTTACGGGCTTGGACCTTACAGCCCGCGGGGGGCTTCGGGGAG 600
|||||
Db 134655 AACCCCGGCATCTTACGGGCTTGGACCTTACAGCCCGCGGGGGGCTTCGGGGAG 134714
|||||
QY 601 AGTCGTAGCGCGCGCGAGCTCTGGGCGCGCCAAAGCGTTTCTGTCTATCCCGGCTACGTG 660
|||||
Db 134715 AGTCGTAGCGCGCGCGAGCTCTGGGCGCGCCAAAGCGTTTCTGTCTATCCCGGCTACGTG 134774
|||||
QY 661 GAGACGCTGTGTGTCGCGGACGAGTCAATGTCTCAAGTTTCCAGCGCGGAGCTTGAACAT 720
|||||
Db 134775 GAGACGCTGTGTGTCGCGGACGAGTCAATGTCTCAAGTTTCCAGCGCGGAGCTTGAACAT 134834
|||||
QY 721 TATCTGTGACGCTGCTGCAACGCGCGGCGGACTCTACGCCATCCAGCATCTCTCAAC 780
|||||
Db 134835 TATCTGTGACGCTGCTGCAACGCGCGGCGGACTCTACGCCATCCAGCATCTCTCAAC 134894
|||||
QY 781 CCATCAACATCTGTTGTGGTCAAGTGTCTTCTTAGAGATCGTGACTCCGGGCGCCAAAG 840
|||||
Db 134895 CCATCAACATCTGTTGTGGTCAAGTGTCTTCTTAGAGATCGTGACTCCGGGCGCCAAAG 134954
|||||
...QY 841 GTCACCGCAATCGCGGCTTACGCTGCGCAACTTCTGTGCTGCGCAGAGAAGCTGAAC 900
|||||
Db 134955 GTCACCGCAATCGCGGCTTACGCTGCGCAACTTCTGTGCTGCGCAGAGAAGCTGAAC 135014
|||||
QY 901 AAGTGTAGTGACAAACACCCCGAGTACTGGGACACTGCATCTCTTCCACAGGACGG 958
|||||
Db 135015 AAGTGTAGTGACAAACACCCCGAGTACTGGGACACTGCATCTCTTCCACAGGACGG 135072
|||||
```

RESULT 5

```
AC025130/c 172905 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 11 clone RP11-211H6 map 11, WORKING DRAFT
DEFINITION SEQUENCE, 16 unordered pieces.
ACCESSION AC025130
VERSION AC025130.2 GI:7387384
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 172905)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-211H6
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 172905)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campoliano,A., Castle,A., Choeil,Y., Colangelo,M., Collins,S.,
Collumore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galegan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
```

Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrim, J., Menues, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL Direct Submission
 Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 172905)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Landers, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, P.,
 Boguslavskiy, L., Bouckghalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrim, J., Menues, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Apr 3, 2000 this sequence version replaced gl:7158941.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7808

Center clone name: 211_H_6

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 161899 bases at least Q40

Consensus quality: 167402 bases at least Q30

Consensus quality: 169801 bases at least Q20

Insert size: 177000; agarose-fp

Insert size: 171405; sum-of-contigs

Quality coverage: 4.1 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1

21: contig of 21 bp in length

* 122 121: gap of 100 bp
 * 122 1343: contig of 1222 bp in length
 * 1344 1443: gap of 100 bp
 * 1444 2636: contig of 1193 bp in length
 * 2637 2736: gap of 100 bp
 * 2737 4052: contig of 1316 bp in length
 * 4053 4152: gap of 100 bp
 * 4153 8044: contig of 3892 bp in length
 * 8045 8144: gap of 100 bp
 * 8145 14949: contig of 6805 bp in length
 * 14950 15049: gap of 100 bp
 * 15050 23535: contig of 8486 bp in length
 * 23536 23635: gap of 100 bp
 * 23636 30261: contig of 6626 bp in length
 * 30262 30361: gap of 100 bp
 * 30362 40941: contig of 10580 bp in length
 * 40942 41041: gap of 100 bp
 * 41042 54893: contig of 13852 bp in length
 * 54894 54993: gap of 100 bp
 * 54994 68888: contig of 13895 bp in length
 * 68889 68988: gap of 100 bp
 * 68989 83702: contig of 14714 bp in length
 * 83703 83802: gap of 100 bp
 * 83803 98393: contig of 14591 bp in length
 * 98394 98493: gap of 100 bp
 * 98494 117356: contig of 18863 bp in length
 * 117357 117456: gap of 100 bp
 * 117457 140589: contig of 23133 bp in length
 * 140590 140689: gap of 100 bp
 * 140690 172905: contig of 32216 bp in length.

FEATURES

source

1..172905
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="11"
 /clone="RP11-211H6"
 /clone_lib="RPC1-11 Human Male BAC"

misc_feature

1..21
 /note="assembly_fragment"
 clone_end:7
 vector_side:right"

misc_feature

122..1343
 /note="assembly_fragment"

misc_feature

1444..2636
 /note="assembly_fragment"

misc_feature

2737..4052
 /note="assembly_fragment"

misc_feature

4153..8044
 /note="assembly_fragment"

misc_feature

8145..14949
 /note="assembly_fragment"

misc_feature

15050..23535
 /note="assembly_fragment"

misc_feature

23636..30261
 /note="assembly_fragment"

misc_feature

30362..40941
 /note="assembly_fragment"

misc_feature

41042..54893
 /note="assembly_fragment"

misc_feature

54994..68888
 /note="assembly_fragment"

misc_feature

68989..83702
 /note="assembly_fragment"

misc_feature

83803..98393
 /note="assembly_fragment"

misc_feature

98494..117356
 /note="assembly_fragment"

misc_feature

117457..140589
 /note="assembly_fragment"

misc_feature

140690..172905

RESULT 6	AC023429	170682 bp	DNA	linear	HTG 10-APR-2001
LOCUS	AC023429	Homo sapiens chromosome 11 clone RP11-121M22, WORKING DRAFT			
DEFINITION	AC023429	SEQUENCE, 7 unordered pieces.			
ACCESSION	AC023429	GI:13569974			
VERSION	AC023429.19	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 170682)			
AUTHORS	Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federpiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hymann, R., Mac, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M., Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D., Southwick, A.M., Webb, C., Wilhelmy, J., Yu, S. and Davis, R.W.				
JOURNAL	Unpublished				
AUTHORS	2	(bases 1 to 170682)			
REFERENCE	Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federpiel, N., Glukhov, S., Hansen, N., Hymann, R., Mac, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelmy, J., Yu, S. and Davis, R.W.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-FEB-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
COMMENT	On Apr 10, 2001 this sequence version replaced gi:13562078.				
	Center: Stanford DNA Sequencing and Technology Development				
	Center code: SDSTDC				
	Web site: http://sequence-www.stanford.edu/group/human/				
	Contact: hum-info@sequence.stanford.edu				
	Project Information				
	Center project name: 837				
	Center clone name: RP11-121M22				
	Summary Statistics				
	Chemistry: Dye-primer; 12% of reads				
	Chemistry: Dye-terminator Big Dye; 86% of reads				
	Assembly program: Phrap; version 0.990319				
	Consensus quality: 165770 bases at least Q40				
	Consensus quality: 167255 bases at least Q30				
	Consensus quality: 167918 bases at least Q20				
	Insert size: 172423; agarose-fp				
	Quality coverage: 9.0x in Q20 bases; agarose-fp				
	Quality coverage: 9.1x in Q20 bases; sum-of-contigs.				
	* NOTE: This is a 'working draft' sequence. It currently				
	* consists of 7 contigs. The true order of the pieces				
	* is not known and their order in this sequence record is				
	* arbitrary. Gaps between the contigs are represented as				
	* runs of N, but the exact sizes of the gaps are unknown.				
	* This record will be updated with the finished sequence				
	* as soon as it is available and the accession number will				
	* be preserved.				
	1	1856: contig of 1856 bp in length			
	1857	1956: gap of unknown length			
	1957	7065: contig of 5109 bp in length			
	7066	7165: gap of unknown length			
	7166	21089: contig of 13924 bp in length			
	21090	21189: gap of unknown length			
	21190	34145: contig of 12956 bp in length			
	34146	34245: gap of unknown length			
	34246	48562: contig of 14317 bp in length			
	48563	48662: gap of unknown length			
	48663	109179: contig of 60517 bp in length			
	109180	109279: gap of unknown length			
	109280	170682: contig of 61403 bp in length.			
FEATURES	Location/Qualifiers				
source	1..170682				

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone_lib="RFLI human BAC library 11"
1..1856
/feature="assembly_name:Contig25"
1957..7065
/feature="assembly_name:Contig26"
7166..21089
/feature="assembly_name:Contig27"
21190..34145
/feature="assembly_name:Contig28"
34246..48562
/feature="assembly_name:Contig29"
48663..109179
/feature="assembly_name:Contig30"
109280..170682
/feature="assembly_name:Contig31"
clone_end:T7"
BASE COUNT 43947 a 39732 c 39514 g 46884 t 605 others
ORIGIN
Query Match 28.1%; Score 803; DB 2; Length 170682;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 156 ACTATTTTTCAGATCACAGCATTTTCAGGAGGACTTTTACCTACACCTGACGCGCGGATGC 215
Db 98615 ACTATTTTTCAGATCACAGCATTTTCAGGAGGACTTTTACCTACACCTGACGCGCGGATGC 98674
QY 216 TCAGTTCTTGGCTCCGCGCTTCTCCATGAGCATCTGGCGCTCCGCGCTCCAGGGGCTCAC 275
Db 98675 TCAGTTCTTGGCTCCGCGCTTCTCCATGAGCATCTGGCGCTCCGCGCTCCAGGGGCTCAC 98734
QY 276 CGGGGGCTTTCAGACCTGCCAGCTGCTTCTATTCGGGGACCTGAACCGCGAGCGGA 335
Db 98735 CGGGGGCTTTCAGACCTGCCAGCTGCTTCTATTCGGGGACCTGAACCGCGAGCGGA 98794
QY 336 CTCGTTCGCTGCTGTGAGCCTGTGCGGGGGCTCCGCGGAGCCTTTGGCTACCGAGCGGC 395
Db 98795 CTCGTTCGCTGCTGTGAGCCTGTGCGGGGGCTCCGCGGAGCCTTTGGCTACCGAGCGGC 98854
QY 396 CGAGTATGCTATAGCCGCTGCCAATGCTAGCGCGCCGCGGCGCGAGCGGCGAAGCGCA 455
Db 98855 CGAGTATGCTATAGCCGCTGCCAATGCTAGCGCGCCGCGGCGCGAGCGGCGAAGCGCA 98914
QY 456 GGGCGCACACCTTCTCCAGCGCGGGGTGTTCGCGGGCGGGCTTCCGGAGACCGCCACTC 515
Db 98915 GGGCGCACACCTTCTCCAGCGCGGGGTGTTCGCGGGCGGGCTTCCGGAGACCGCCACTC 98974
QY 516 TCCTGCGGGGTGGCTCGGGCTGGAAACCCCGCATCTACGGGCGGCTTGGAGCCTTACAA 575
Db 98975 TCCTGCGGGGTGGCTCGGGCTGGAAACCCCGCATCTACGGGCGGCTTGGAGCCTTACAA 99034
QY 576 CCGCGGGGGCGGGCTTCCGGGGAGAGTCTGTAGCGCGCGAGGCTGGGGCGCGCAAGCG 635
Db 99035 CCGCGGGGGCGGGCTTCCGGGGAGAGTCTGTAGCGCGCGAGGCTGGGGCGCGCAAGCG 99094
QY 636 TTTCGTGCTATCCCGGGTACGTGGAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 695
Db 99095 TTTCGTGCTATCCCGGGTACGTGGAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 99154
QY 696 GTTCCAGGGGGCGGACCTGGAACATTATCTGTGACGCTGCTGCGACGCGCGCGGCGACT 755
Db 99155 GTTCCAGGGGGCGGACCTGGAACATTATCTGTGACGCTGCTGCGACGCGCGGCGGACT 99214
QY 756 CTACCGGCATCCAGCATCCTCAACCCCATCAACATGTTGTTGTTGTTGTTGTTGTTGTTGTT 815
Db 99215 CTACCGGCATCCAGCATCCTCAACCCCATCAACATGTTGTTGTTGTTGTTGTTGTTGTTGTT 99274
QY 816 TAGAGATCGTGACTCCCGGGGGCGGAGGTACCGGCAATGCGGGCCTGACGCTGCGCAACTT 875
```

```
Db 99275 TAGAGATCGTGACTCCCGGGGGCGGAGGTACCGGCAATGCGGGCCTGACGCTGCGCAACTT 99334
QY 876 CTGTGCTGGCAGAGAAGCTGAACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 935
Db 99335 CTGTGCTGGCAGAGAAGCTGAACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 99394
QY 936 TGCCATCCTCTTCCACGAGGAGG 958
Db 99395 TGCCATCCTCTTCCACGAGGAGG 99417

RESULT 7
LOCUS AP003459 28000 bp DNA linear PRI 08-MAR-2002
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-211H5,
complete sequence.
ACCESSION AP003459
VERSION AP003459.2 GI:19263045
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-211H5.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 28000)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel.81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Mar 7, 2002 this sequence version replaced gi:13488920.
FEATURES
Source
1..28000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-211H5"
BASE COUNT 6938 a 6083 c 6679 g 8300 t
ORIGIN
Query Match 25.4%; Score 724; DB 9; Length 28000;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2079 GCATGCTACAATTCGTGTGGCCATCCCGCAGCGGCTCAAGCATCGACATCGCCA 2138
Db 3473 GCATGCTACAATTCGTGTGGCCATCCCGCAGCGGCTCAAGCATCGACATCGCCA 3532
QY 2139 GCGCGTTTACAAGGGCTCATCGGGATGACAACCTACCTGGCTCTGAAGAACAGGCAAGG 2198
Db 3533 GCGCGTTTACAAGGGCTCATCGGGATGACAACCTACCTGGCTCTGAAGAACAGGCAAGG 3592
QY 2199 CAAGTACCTGCTCAAGGGCATTTCTGTGTGTCGGGTGGAGCGGACCTGGTGTGAA 2258
Db 3593 CAAGTACCTGCTCAAGGGCATTTCTGTGTGTCGGGTGGAGCGGACCTGGTGTGAA 3652
QY 2259 GGGCAGTCTGCTGCGGTACAGCGGACGCGGACAGCGGTGGAGAGCTTCAGGCTTCCCG 2318
Db 3653 GGGCAGTCTGCTGCGGTACAGCGGACGCGGACAGCGGTGGAGAGCTTCAGGCTTCCCG 3712
QY 2319 GCGCATCTGGAGCGCTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACGCGCCCG 2378
Db 3713 GCGCATCTGGAGCGCTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACGCGCCCG 3772
```



```

Qy 2379 GGTCCGCTACTCTTCTATCTGCCAAAGAGCTCGGAGGACAACTCTCTCATCCCAA 2438
Db 3773 GGTCCGCTACTCTTCTATCTGCCAAAGAGCTCGGAGGACAACTCTCTCATCCCAA 3832
Qy 2439 GGACCCCGGGGACCCCTCTCTTGACACACGCTCCTCAGCCTCTCCAAACAGGTGA 2498
Db 3833 GGACCCCGGGGACCCCTCTCTTGACACACGCTCCTCAGCCTCTCCAAACAGGTGA 3892
Qy 2499 GCAGCCGGAGCAGAGCCGCCCTCAGCCTGGGTGGCTGGCAGCTGGGGGCGCTGCTCCCG 2558
Db 3993 GCAGCCGGAGCAGAGCCGCCCTCAGCCTGGGTGGCTGGCAGCTGGGGGCGCTGCTCCCG 3952
Qy 2559 GAGTCGGCGAGTGGCTCGCAGAAGCGGGCGGTGGACTGTGCGGGCTCGCGCGGAGCGG 2618
Db 3953 GAGTCGGCGAGTGGCTCGCAGAAGCGGGCGGTGGACTGTGCGGGCTCGCGCGGAGCGG 4012
Qy 2619 CACGGTCCCTGCTGTGATCAGCCATCGGCGCGTGGAGACACAAAGCTTGGGGGAGCC 2678
Db 4013 CACGGTCCCTGCTGTGATCAGCCATCGGCGCGTGGAGACACAAAGCTTGGGGGAGCC 4072
Qy 2679 CTGCCCCACCTGGGAGCTCAGCCCTGGTCAACCTGCTCCTCAGAGCTGGCGCGCGGGATT 2738
Db 4073 CTGCCCCACCTGGGAGCTCAGCCCTGGTCAACCTGCTCCTCAGAGCTGGCGCGCGGGATT 4132
Qy 2739 TCAGAGCGCTCACTCAAGTGTGTGGCCACGAGCGCGCTCTGCGCGGAGCCAGTG 2798
Db 4133 TCAGAGCGCTCACTCAAGTGTGTGGCCACGAGCGCGCTCTGCGCGGAGCCAGTG 4192
Qy 2799 CAACCTGCACCGCAAGCCCGCAGGAGCTGGACTTCTGCTGCTCAGGCGCGTGTGTA 2853
Db 4193 CAACCTGCACCGCAAGCCCGCAGGAGCTGGACTTCTGCTGCTCAGGCGCGTGTGTA 4247

RESULT 8
AC025130
LOCUS Homo sapiens chromosome 11 clone RP11-211H6 map 11, WORKING DRAFT
DEFINITION
SEQUENCE, 16 unordered pieces.
AC025130
VERSION AC025130.2 GI:7387384
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 172905)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, G., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
Direct Submission
TITLE
JOURNAL

```

REFERENCE AUTHORS

Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 172905)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
 Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, G., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 3, 2000 this sequence version replaced gi:7138941.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L7808
 Center clone name: 211_H_6
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 161899 bases at least Q40
 Consensus quality: 167402 bases at least Q30
 Consensus quality: 169801 bases at least Q20
 Insert size: 17700; agarose-
 Quality coverage: 4.1 in Q20 bases; agarose-
 Quality coverage: 4.2 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1
 * 22 121: gap of 100 bp
 * 122 1343: contig of 1222 bp in length
 * 1344 1443: gap of 100 bp
 * 1444 2636: contig of 1193 bp in length
 * 2637 2736: gap of 100 bp
 * 2737 4052: contig of 1316 bp in length
 * 4053 4152: gap of 100 bp
 * 4153 8044: contig of 3892 bp in length
 * 8045 8144: gap of 100 bp
 * 8145 14949: contig of 6805 bp in length
 * 14950 15049: gap of 100 bp
 * 15050 23535: contig of 8486 bp in length
 * 23536 23636: gap of 100 bp
 * 23636 30261: contig of 6626 bp in length


```

* 30262 30361: gap of 100 bp
* 30362 40941: contig of 10580 bp in length
* 40942 41041: gap of 100 bp
* 41042 54893: contig of 13852 bp in length
* 54894 54993: gap of 100 bp
* 54994 68888: contig of 13895 bp in length
* 68889 69888: gap of 100 bp
* 69889 83702: contig of 14714 bp in length
* 83703 83802: gap of 100 bp
* 83803 98393: contig of 14591 bp in length
* 98394 98493: gap of 100 bp
* 98494 117356: contig of 18863 bp in length
* 117357 117456: gap of 100 bp
* 117457 140589: contig of 23133 bp in length
* 140590 140689: gap of 100 bp
* 140690 172905: contig of 32216 bp in length.

```

FEATURES

Source

```

1. 172905
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="11"
/clone_lib="RP11-211H6"
/clone_lib="RPC1-11 Human Male BAC"
1. 21

```

misc_feature

```

/note="assembly_fragment
clone_end:17
vector_side:right"

```

misc_feature

```

/note="assembly_fragment"
122..1343

```

misc_feature

```

/note="assembly_fragment"
1444..2636

```

misc_feature

```

/note="assembly_fragment"
2737..4052

```

misc_feature

```

/note="assembly_fragment"
4153..8044

```

misc_feature

```

/note="assembly_fragment"
8145..14949

```

misc_feature

```

/note="assembly_fragment"
15050..23535

```

misc_feature

```

/note="assembly_fragment"
23636..30261

```

misc_feature

```

/note="assembly_fragment
clone_end:SP6
vector_side:right"

```

misc_feature

```

/note="assembly_fragment"
30362..40941

```

misc_feature

```

/note="assembly_fragment"
41042..54893

```

misc_feature

```

/note="assembly_fragment"
54994..68888

```

misc_feature

```

/note="assembly_fragment"
68989..83702

```

misc_feature

```

/note="assembly_fragment"
83803..98393

```

misc_feature

```

/note="assembly_fragment"
98494..117356

```

misc_feature

```

/note="assembly_fragment"
117457..140589

```

misc_feature

```

/note="assembly_fragment"
140690..172905

```

misc_feature

```

/note="assembly_fragment"
44360 a 41333 c 39516 g 46194 t 1502 others

```

BASE COUNT

```

ORIGIN
Query Match 12.5%; Score 358; DB 2; Length 172905;

```

```

Best Local Similarity 100.0%; Pred. No. 1.4e-184;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2079 GCATGGCTACAAATTCGTGGTGCCATCCCGCAGCGCCTCAAGCATCGACATCCGCCA 2138
|||||

```

```

Db 54519 GCATGGCTACAAATTCGTGGTGCCATCCCGCAGCGCCTCAAGCATCGACATCCGCCA 54578
|||||

```

```

Qy 2139 GCGCGGTTACAAAGGCGTATCGGGGATGACAACACTACCTGGCTCTGAAGAACAGCAAGG 2198
|||||

```

```

Db 54579 GCGCGGTTACAAAGGCGTATCGGGGATGACAACACTACCTGGCTCTGAAGAACAJCCCAAGG 54638
Qy 2199 CAAGTACCTGCTCAACAGGGCATTTCTGTTGTCGGCGGTGAGCGGGACCTGGTGGTCAA 2258
|||||
Db 54639 CAAGTACCTGCTCAACAGGGCATTTCTGTTGTCGGCGGTGAGCGGGACCTGGTGGTCAA 54698
Qy 2259 GGGCAGTCTGCTGCGGTACAGCGGCACAGCGGTGAGAGCCTGAGAGCTTCCG 2318
|||||
Db 54699 GGGCAGTCTGCTGCGGTACAGCGGCACAGCGGTGAGAGCCTGAGAGCTTCCG 54758
Qy 2319 GCCATCTCTGGAGCGCTGACCGTGGAGTCTCTCCCTGGGGAAGATGACACGCCGCCG 2378
|||||
Db 54759 GCCATCTCTGGAGCGCTGACCGTGGAGTCTCTCCCTGGGGAAGATGACACGCCGCCG 54818
Qy 2379 GTTCCGCTACTCTCTTATCTATCTGCCCCAAGAGCCTCGGAGGACCAAGTCTCTCATCC 2436
|||||
Db 54819 GGTCCGCTACTCTCTTATCTATCTGCCCCAAGAGCCTCGGAGGACCAAGTCTCTCATCC 54876

```

RESULT 9

AC023429/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center

Center code: SBDTDC

Web site: http://sequence-www.stanford.edu/group/human/

Contact: hum-info@sequence.stanford.edu

----- Project Information

Center project name: 837

Center clone name: RP11-121M22

----- Summary Statistics

Sequencing Vector: M13mp18; X02513

Chemistry: Dye-primer; 12% of reads

Chemistry: Dye-terminator Big Dye; 86% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 165770 bases at least Q40

Consensus quality: 167255 bases at least Q30

Consensus quality: 167918 bases at least Q20

Insert size: 172423; agarose-fp

Insert size: 170082; sum-of-contigs

Quality coverage: 9.0x in Q20 bases; agarose-fp

Quality coverage: 9.1x in Q20 bases; sum-of-contigs.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 7 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

```

170682 bp DNA linear HT: 10-APR-2001
Homo sapiens chromosome 11 clone RP11-121M22, WORKING DRAFT
SEQUENCE, 7 unordered pieces.

```

```

AC023429 AC023429.19 GI:13569974
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Homo sapiens
Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; H.omo.
1 (bases 1 to 170682)
Abola, A. P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D.,

```

```

Fedeispiet, N., Glukhov, S., Hansen, N., Herman, Z. S., Hymen, R.,
Mao, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M.,
Morehouse, A. J., Nguyen, M., Oefner, P., Palm, C. J., Ramirez, D.,
Southwick, A. M., Webb, C., Wilhelmy, J., Yu, S. and Davis, R. W.
Unpublished
2 (bases 1 to 170682)

```

```

Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Fedeispiet, N.,
Glukhov, S., Hansen, N., Hymen, R., Mao, J., Marathe, R.,
Morehouse, A. J., Oefner, P., Palm, C. J., Ramirez, D., Wilhelmy, J.,
Yu, S. and Davis, R. W.
Direct Submission
Submitted (14-FEB-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

```

```

On Apr 10, 2001 this sequence version replaced gi:13562078.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development
Center

```

```

Center code: SBDTDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: 837
Center clone name: RP11-121M22
----- Summary Statistics
Sequencing Vector: M13mp18; X02513
Chemistry: Dye-primer; 12% of reads
Chemistry: Dye-terminator Big Dye; 86% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165770 bases at least Q40
Consensus quality: 167255 bases at least Q30
Consensus quality: 167918 bases at least Q20
Insert size: 172423; agarose-fp
Insert size: 170082; sum-of-contigs
Quality coverage: 9.0x in Q20 bases; agarose-fp
Quality coverage: 9.1x in Q20 bases; sum-of-contigs.

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1856: contig of 1856 bp in length
 * 1957 1956: gap of unknown length
 * 1957 7085: contig of 5109 bp in length
 * 7066 7165: gap of unknown length
 * 7166 21089: contig of 13924 bp in length
 * 21090 21189: gap of unknown length
 * 21190 34145: contig of 12956 bp in length
 * 34146 34245: gap of unknown length
 * 34246 48562: contig of 14317 bp in length
 * 48563 48662: gap of unknown length
 * 48663 109179: contig of 60517 bp in length
 * 109180 109279: gap of unknown length
 * 109280 170682: contig of 61403 bp in length.

FEATURES

source
 1. .170682
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /clone="RP11-121M22"
 /clone_lib="RPCI human BAC library 11"
 1. 1856
 /note="assembly_name:Contig25"
 1957. 7065
 /note="assembly_name:Contig26"
 7166. 21089
 /note="assembly_name:Contig27"
 21190. 34145
 /note="assembly_name:Contig28"
 34246. 48562
 /note="assembly_name:Contig29
 clone_end:SP6"
 48663. 109179
 /note="assembly_name:Contig30"
 109280. 170682
 /note="assembly_name:Contig31
 clone_end:T7"
 BASE COUNT 43947 a 39732 c 39514 g 46884 t 605 others
 ORIGIN
 Query Match 10.1%; Score 288; DB 2: Length 170682;
 Best Local Similarity 100.0%; Pred. No. 3.9e-146;
 Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1258 GGTGACTGCTCTGGACCAACCCAGCAAGCCATCTCCCTGCGCCGAGGATCTGCCGGGC 1317
 |||||
 DB 44571 GGTGACTGCTCTGGACCAACCCAGCAAGCCATCTCCCTGCGCCGAGGATCTGCCGGGC 44512
 |||||
 QY 1318 GCCAGCTACACCTGAGCCAGCAGTGGAGTGGCTTTTGGCGTGGGCTTCAAGCCCTGT 1377
 |||||
 DB 44511 GCCAGCTACACCTGAGCCAGCAGTGGAGTGGCTTTTGGCGTGGGCTTCAAGCCCTGT 44452
 |||||
 QY 1378 CCTTACATGCACTGTCACCAAGCTGTGTGTCACCGGAGGCAAGGACAGATGCTG 1437
 |||||
 DB 44451 CCTTACATGCACTGTCACCAAGCTGTGTGTCACCGGAGGCAAGGACAGATGCTG 44392
 |||||
 QY 1438 TGCCAGACCCGCACTTCCCTGGGCGGATGGCCACCAAGCTGTGGCGAGGCAAGCTCTGC 1497
 |||||
 DB 44391 TGCCAGACCCGCACTTCCCTGGGCGGATGGCCACCAAGCTGTGGCGAGGCAAGCTCTGC 44332
 |||||
 QY 1498 CTCRAAGGGGCTGGTGGAGAGACACACCTCAACAGACACAGGCTG 1545
 |||||
 DB 44331 CTCRAAGGGGCTGGTGGAGAGACACACCTCAACAGACACAGGCTG 44284
 |||||

RESULT 10

AC101990 182656 bp DNA linear HTG 21-AUG-2002
 LOCUS Mus musculus clone RP24-371J2, WORKING DRAFT SEQUENCE, 16 unordered
 DEFINITION pieces.

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AC101990
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 house mouse.
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 182656)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP24-371J2

REFERENCE AUTHORS TITLE JOURNAL

Unpublished
 2 (bases 1 to 182656)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Melgrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL

Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE AUTHORS

3 (bases 1 to 182656)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Melgrim,J., Meneus,L., Mihova,T., Norbu,C., Norman,C.H.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Peterson,K.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 21, 2002 this sequence version replaced gi:17060766.
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L17833
 Center clone name: 371_J-2
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 178030 bases at least Q40
 Consensus quality: 179964 bases at least Q30
 Consensus quality: 180647 bases at least Q20
 Insert size: 172000; agarose-1p
 Insert size: 181156; sum-of-contigs
 Quality coverage: 8.1 in Q20 bases; agarose-1p
 Quality coverage: 7.7 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      361: contig of 361 bp in length
*      362 461: gap of 100 bp
*      462 1538: contig of 1077 bp in length
*      1539 1638: gap of 100 bp
*      1639 3397: contig of 1759 bp in length
*      3398 3497: gap of 100 bp
*      3498 4947: contig of 1450 bp in length
*      4948 5047: gap of 100 bp
*      5048 6893: contig of 1846 bp in length
*      6894 6993: gap of 100 bp
*      6994 10395: contig of 3402 bp in length
*      10396 10495: gap of 100 bp
*      10496 12584: contig of 2089 bp in length
*      12585 12684: gap of 100 bp
*      12685 14977: contig of 2293 bp in length
*      14978 15077: gap of 100 bp
*      15078 18181: contig of 3104 bp in length
*      18182 18281: gap of 100 bp
*      18282 22654: contig of 4373 bp in length
*      22655 22754: gap of 100 bp
*      22755 31261: contig of 8507 bp in length
*      31262 31361: gap of 100 bp
*      31362 43587: contig of 12226 bp in length
*      43588 43687: gap of 100 bp
*      43688 57612: contig of 13925 bp in length
*      57613 57712: gap of 100 bp
*      57713 77467: contig of 19755 bp in length
*      77468 77567: gap of 100 bp
*      77568 108724: contig of 31157 bp in length
*      108725 108824: gap of 100 bp
*      108825 182656: contig of 73832 bp in length.

```

FEATURES

```

source
1. .182656
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-371J2"
/clone_lib="RPC1-24 Male Mouse BAC"
misc_feature
1. 361
/notes="assembly_fragment"
462. 1538
/notes="assembly_fragment"
1639. 3397
/notes="assembly_fragment"
3498. 4947
/notes="assembly_fragment"
5048. 6893
/notes="assembly_fragment"
6994. 10395
/notes="assembly_fragment"
10496. 12584
/notes="assembly_fragment"
12685. 14977
/notes="assembly_fragment"
15078. 18181
/notes="assembly_fragment"

```

```

misc_feature 18282..22654
/notes="assembly_fragment"
misc_feature 22755..31261
/notes="assembly_fragment"
misc_feature 31362..43587
/notes="assembly_fragment"
misc_feature 43688..57612
/notes="assembly_fragment"
misc_feature 57713..77467
/notes="assembly_fragment"
misc_feature 77568..108724
/notes="assembly_fragment"
misc_feature 108825..182656
/notes="assembly_fragment"

BASE COUNT 51468 a 41160 c 39770 g 48758 t 1500 others
ORIGIN

Query Match 2.6%: Score 74; DB 2: Length 182656;
Best Local Similarity 100.0%; Pred. No. 1.4e-28;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1009 ACCATGTGTGACCCCAAGAGAGAGCTGCTGTCTCATTGAGGACGATGGCTTCCATCAGCC 1068
|||||
Db 16446 ACCATGTGTGACCCCAAGAGAGAGCTGCTGTCTCATTGAGGACGATGGCTTCCATCAGCC 16505
|||||

QY 1069 TTCACCACCTGCCCA 1082
|||||
Db 16506 TTCACCACCTGCCCA 16519
|||||

RESULT 11
AC126507/c
LOCUS
DEFINITION
AC126507
VERSION
AC126507.1 GI:21700463
KEYWORDS
HTG: HTGS_PHASE1.
SOURCE
Rattus norvegicus.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 178764)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Brown,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssohn,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,N., Neal,D., Newtonson,N., Newtonson,N., Nguyen,N., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

```

```

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubohan, I., Roife, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 178764)
Worley, K.C.
Direct Submission
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178764)
Worley, K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G2FU
Center clone name: CH230-254N12
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124573 bases at least Q40
Consensus quality: 131850 bases at least Q30
Consensus quality: 137700 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1037: contig of 1037 bp in length
* 1038: gap of unknown length
* 1138: contig of 1113 bp in length
* 2251: gap of unknown length
* 2350: gap of unknown length
* 2351: contig of 1229 bp in length
* 3580: gap of unknown length
* 3680: contig of 1195 bp in length
* 4874: gap of unknown length
* 4875: contig of 1551 bp in length
* 4975: gap of unknown length
* 6526: gap of unknown length
* 6626: contig of 1280 bp in length
* 7906: gap of unknown length
* 8006: contig of 1719 bp in length
* 9824: gap of unknown length
* 9825: contig of 1618 bp in length
* 11443: gap of unknown length
* 11543: contig of 1358 bp in length
* 12901: gap of unknown length
* 13001: contig of 1150 bp in length
* 14151: gap of unknown length
* 14251: contig of 1315 bp in length
* 15565: gap of unknown length
* 15566: contig of 1037 bp in length
*
* 15656: contig of 1810 bp in length
* 17476: gap of unknown length
* 18690: contig of 1115 bp in length
* 18790: gap of unknown length
* 20269: contig of 1479 bp in length
* 20363: gap of unknown length
* 21927: contig of 1558 bp in length
* 22027: gap of unknown length
* 22028: contig of 1971 bp in length
* 23998: gap of unknown length
* 25679: contig of 1581 bp in length
* 25773: gap of unknown length
* 27428: contig of 1649 bp in length
* 27528: gap of unknown length
* 29677: contig of 2149 bp in length
* 29777: gap of unknown length
* 30877: contig of 1100 bp in length
* 30977: gap of unknown length
* 32831: contig of 1854 bp in length
* 32931: gap of unknown length
* 35419: contig of 2488 bp in length
* 35519: gap of unknown length
* 37393: contig of 1874 bp in length
* 37493: gap of unknown length
* 38335: contig of 2342 bp in length
* 39335: gap of unknown length
* 42789: contig of 2854 bp in length
* 42889: gap of unknown length
* 46646: contig of 3757 bp in length
* 46746: gap of unknown length
* 49188: contig of 2442 bp in length
* 49288: gap of unknown length
* 53805: contig of 4517 bp in length
* 53905: gap of unknown length
* 56202: contig of 2297 bp in length
* 56302: gap of unknown length
* 58719: contig of 2417 bp in length
* 58819: gap of unknown length
* 62904: contig of 4085 bp in length
* 63004: gap of unknown length
* 63005: contig of 2928 bp in length
* 65932: gap of unknown length
* 65933: gap of unknown length
* 70120: contig of 4088 bp in length
* 70220: gap of unknown length
* 73258: contig of 3038 bp in length
* 73358: gap of unknown length
* 77466: contig of 4108 bp in length
* 77566: gap of unknown length
* 80545: contig of 2979 bp in length
* 80645: gap of unknown length
* 85987: contig of 5342 bp in length
* 86087: gap of unknown length
* 86088: contig of 4821 bp in length
* 90909: gap of unknown length
* 91009: contig of 6910 bp in length
* 97919: gap of unknown length
* 98019: contig of 5484 bp in length
* 103502: gap of unknown length
* 103503: contig of 8968 bp in length
* 112570: gap of unknown length
* 112571: contig of 5254 bp in length
* 117924: gap of unknown length
* 118024: contig of 6995 bp in length
* 125019: gap of unknown length
* 125020: contig of 8968 bp in length
* 134088: gap of unknown length
* 134188: contig of 9664 bp in length
* 143852: gap of unknown length
* 143952: contig of 7748 bp in length
* 151700: gap of unknown length
* 151799: contig of 7027 bp in length
* 158826: gap of unknown length
* 158927: contig of 9596 bp in length
* 168522: contig of 9596 bp in length

```

```

* 168523 168622: gap of unknown length
* 168623 178764: contig of 10142 bp in length.
FEATURES
    Location/Qualifiers
        1..178764
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
            /clone="CH230-254N12"
BASE COUNT 47487 a 39119 c 38897 g 45423 t 7838 others
ORIGIN
    Query Match
    Best Local Similarity 100.08; Pred. No. 2e-21;
    Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1942 GTCTGTGTCACGAAGTCATCAAGCTGCTGTGTGATGGGAACCTGGGCTCCAAGAAG 2001
    |||||
Db 150003 GTCTGTGTCACGAAGTCATCAAGCTGCTGTGTGATGGGAACCTGGGCTCCAAGAAG 149944
QY 2002 A 2002
Db 149943 A 149943

RESULT 12
AC126507
LOCUS      178764 bp      DNA      linear      HTG 24-JUL-2002
DEFINITION Rattus norvegicus clone CH230-254N12, *** SEQUENCING IN PROGRESS
AC126507
VERSION    AC126507.1 GI:21700463
KEYWORDS   HTG; HTGS-PHASE1.
SOURCE     Rattus norvegicus.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE
    1 (bases 1 to 178764)
        Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
        Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
        Barbara, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
        Boucek, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
        Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
        Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
        Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
        Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
        Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
        Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
        Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
        Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
        Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
        Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
        Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
        Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
        Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
        Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
        Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
        Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
        Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
        Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulised, H.,
        Lozadro, R.J., Lu, X., Lucier, A., Lucier, R., Luna, P., Ma, J.,
        Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
        Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
        Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
        Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
        Nguyen, N., Nickerson, E., Nwokenwo, S., Ogwu, M., Okwuonu, G.,
        Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
        Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
        Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
        Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
        Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
        Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
        Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
        Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,

```

Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 178764)
Worley, K.C.
Direct Submission
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178764)
Worley, K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZFU
Center clone name: CH230-254N12
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124573 bases at least Q40
Consensus quality: 131850 bases at least Q30
Consensus quality: 137700 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drafi_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1037: contig of 1037 bp in length
1038 1137: gap of unknown length
1138 2250: contig of 1113 bp in length
2251 2350: gap of unknown length
2351 3579: contig of 1229 bp in length
3580 3679: gap of unknown length
3680 4874: contig of 1195 bp in length
4875 4974: gap of unknown length
4975 6525: contig of 1551 bp in length
6526 6625: gap of unknown length
6626 7905: contig of 1280 bp in length
7906 8005: gap of unknown length
8006 8724: contig of 1719 bp in length
8725 9824: gap of unknown length
9825 11442: contig of 1618 bp in length
11443 12900: contig of 1358 bp in length
12901 13000: gap of unknown length
13001 14150: contig of 1150 bp in length
14151 14250: gap of unknown length
14251 15365: contig of 1315 bp in length
15366 15665: gap of unknown length
15666 17475: contig of 1810 bp in length
17476 17575: gap of unknown length
17576 18690: contig of 1115 bp in length
18691 18790: gap of unknown length
18791 20269: contig of 1479 bp in length
20270 20369: gap of unknown length
20370 21927: contig of 1558 bp in length

* 21928 22027: gap of unknown length
* 22028 23998: contig of 1971 bp in length
* 23999 24098: gap of unknown length
* 24099 25679: contig of 1581 bp in length
* 25680 27428: gap of unknown length
* 27429 27528: gap of unknown length
* 27529 29677: contig of 2149 bp in length
* 29678 29777: gap of unknown length
* 29778 30877: contig of 1100 bp in length
* 30878 30977: gap of unknown length
* 30978 32831: contig of 1854 bp in length
* 32832 32931: gap of unknown length
* 32932 35419: contig of 2488 bp in length
* 35420 35519: gap of unknown length
* 35520 37393: contig of 1874 bp in length
* 37394 37493: gap of unknown length
* 37494 39835: contig of 2342 bp in length
* 39836 39935: gap of unknown length
* 39936 42789: contig of 2854 bp in length
* 42790 42889: gap of unknown length
* 42890 46646: contig of 3757 bp in length
* 46647 46746: gap of unknown length
* 46747 49188: contig of 2442 bp in length
* 49189 49288: gap of unknown length
* 49289 53805: contig of 4517 bp in length
* 53806 53905: gap of unknown length
* 53906 56202: contig of 2297 bp in length
* 56203 56302: gap of unknown length
* 56303 58719: contig of 2417 bp in length
* 58720 58819: gap of unknown length
* 58820 62904: contig of 4085 bp in length
* 62905 63004: gap of unknown length
* 63006 65932: contig of 2928 bp in length
* 65933 66032: gap of unknown length
* 66033 70120: contig of 4088 bp in length
* 70121 70220: gap of unknown length
* 70221 73258: contig of 3038 bp in length
* 73259 73358: gap of unknown length
* 73359 77466: contig of 4108 bp in length
* 77467 77566: gap of unknown length
* 77567 80545: contig of 2979 bp in length
* 80546 80646: gap of unknown length
* 80646 85987: contig of 5342 bp in length
* 85988 86087: gap of unknown length
* 86088 90908: contig of 4821 bp in length
* 90909 91008: gap of unknown length
* 91009 97918: contig of 6910 bp in length
* 97919 98018: gap of unknown length
* 98019 103502: contig of 5484 bp in length
* 103503 103602: gap of unknown length
* 103603 112570: contig of 8968 bp in length
* 112571 112670: gap of unknown length
* 112671 117924: contig of 5254 bp in length
* 117925 118024: gap of unknown length
* 118025 125013: contig of 6995 bp in length
* 125014 125119: gap of unknown length
* 125120 134087: contig of 8968 bp in length
* 134088 134187: gap of unknown length
* 134188 143851: contig of 9664 bp in length
* 143852 143951: gap of unknown length
* 143952 151699: contig of 7748 bp in length
* 151700 151799: gap of unknown length
* 151800 158826: contig of 7027 bp in length
* 158827 158927: gap of unknown length
* 158928 168522: contig of 9596 bp in length
* 168523 178764: contig of 10142 bp in length.
* 186623 Location/Qualifiers
1. 1178764
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-254N12"

FEATURES
source

BASE COUNT 47487 a 39119 c 38897 g 45423 t 7838 others
ORIGIN
Query Match 1.58; Score 44; DB 2; Length 178764;
Best Local Similarity 100.0%; Pred. No. 4.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 511 ACCTCTCGCTCGGGGTGCGCTCGGCTGGAACCCCGCATCCT 554
|||||
Db 101958 ACCTCTCGCTCGGGGTGCGCTCGGCTGGAACCCCGCATCCT 102001
|||||
...
RESULT 13
AC101990/c
LOCUS
DEFINITION Mus musculus clone RP24-371J2, WORKING DRAFT SEQUENCE, 16 unordered
pieces.
AC101990
AC101990.2 GI:22381363
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS house mouse.
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 182656)
Birren,B., Nusbaum,C. and Lander,E.
TITLE
Mus musculus, clone RP24-371J2
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 182656)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gadyana,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Melidrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 182656)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepe,Y., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gadyana,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Melidrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 8 Row: j Column: 3

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

Location/Qualifiers

1..3927

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="CZECH II"

/clone="IMAGE:3491991"

/tissue_type="Mammary tumor metastatized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."

/clone_lib="NCI CGAP_Lu29"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

<1..1024

/codon_start=2

/product="Similar to a disintegrin and metalloproteinase with thrombospondin motifs 1 (ADAMTS-1)"

/protein_id="AAH09667.1"

/db_xref="GI:16307156"

/translation="PRVRKCLICRANGTYFYVLAPKVDGTLCTPDSTVCYQKCIKAGCDNGLGSKKFDKCGCGDNKSKRVYGLTKPMHYNFVVAIPAGASSIDIRQRGYGLIGDDNYLALKNSQGYLLNGHFVSAVERDLVVKGSVLYRSGTGTAVESIQASRPILFLEPTLVSVGKMTPPRVYSFLPKPEPREDKSTRPKDPRGSLVLRNSVLSLSNOVEOPDNRPAPRWAGSWGSPCVSCGSLQKRAVDCRDSPGQGSACDVRHPLFKRACGPEPTWELGNWSPCSKSCGRGFKRRPLKCVGHGRLRLARQCDLRRKPQELDFCVLRPC"

CDS

BASE COUNT 938 a 920 c 1093 g 976 t

ORIGIN

Query Match 1.2%; Score 35; DB 10; Length 3927;

Best Local Similarity 100.0%; Pred. No. 5.3e-07;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1942 GTCTGTCTCCAGGCAAGTCATCAGGCTGGCTG 1976

|||||

Db 113 GTCTGTCTCCAGGCAAGTCATCAGGCTGGCTG 147

Search completed: May 16, 2003, 03:53:15

Job time : 10280 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 22:55:17 ; Search time 577 Seconds
(without alignments)
11135.097 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtctctgtggcctcct.....gcgtctctgaggccgtgtctga 2853

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 24

Total number of hits satisfying chosen parameters: 19

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2853	100.0	2853	24	Human protease cDN
2	2853	100.0	3446	24	Human protease cDN
3	2695	94.5	2853	22	Human metalloprote
4	2196	77.0	2930	24	Human metalloprote
5	2196	77.0	2937	24	Human metalloprote
6	1091	38.2	1104	24	Human protease cDN
7	959	33.6	966	24	Human protease cDN
8	481	16.9	1143	21	Human metalloprote
9	50	1.8	1518	21	Rat metalloprotein

c 10 30 1.1 30 22 AAH41029
c 11 30 1.1 30 22 AAH41030
c 12 30 1.1 41 22 AAH41017
c 13 27 0.9 27 22 AAH41018
c 14 27 0.9 37 22 AAH41019
c 15 27 0.9 38 22 AAH41024
c 16 27 0.9 38 22 AAH41033
c 17 27 0.9 38 22 AAH41034
c 18 25 0.9 610 24 ABQ44966
c 19 25 0.9 610 24 ABQ44967

Adaptor primer SEQ
Adaptor primer SEQ
PCR primer specific
PCR primer specific
PCR primer specific
PCR primer specific
PCR primer for met
PCR primer for met
Oligonucleotide fo
Oligonucleotide fo

ALIGNMENTS

RESULT 1

AAD35569

ID AAD35569 standard; cDNA; 2853 BP.

XX

AC AAD35569;

XX

DT 26-JUL-2002 (first entry)

XX

DE Human protease cDNA #2.

XX

KW Human; novel human protein; NHP; protease; biological disorder; obesity;
KW high blood pressure; arthritis; connective tissue disorder; infertility;
KW gene therapy; enzyme; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..2853

FT

FT /*tag= a

FT /product= "Human protease #2"

PN WO200226949-A2.

PD

PD 04-APR-2002.

PF

PF 27-SEP-2001; 2001WO-US30350.

XX

XX 29-SEP-2000; 2000US-236689P.

PA (LEXI-) LEXICON GENETICS INC.

PI Friddle CJ, Hilbun E;

DR WPI: 2002-372123/40.

XX P-PSDB; AAD22541.

XX

PT Novel nucleic acid encoding a human protease, useful as a hybridization
PT probe for screening libraries and assessing gene expression patterns -
XX Claim 1; Page 35-36; 41pp; English.

XX The present sequence is a cDNA encoding novel human protein (NHP),
CC human protease. NHPs share structural similarity with animal proteases
CC particularly zinc metalloproteases. Sequences of the invention are
CC useful in therapeutic, diagnostic and pharmacogenomic applications.
CC NHP polynucleotides are used as hybridisation probes for screening
CC libraries and assessing gene expression patterns. They can also be
CC used for treating related biological disorders such as obesity, high
CC blood pressure, arthritis, connective tissue disorders and infertility.
CC They are also used in gene therapy.

XX Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other;

XX Query Match 100.0%; Score 2853; DB 24; Length 2853;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGCATCTTAACCTGGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60

QY 2221 TTCGTGGTCTCGCGGTGAGCGGACCTGTGTGTAAG36CAGTCTGCTGCGGTACAGC 2280
Db 2221 TTCGTGGTCTCGCGGTGAGCGGACCTGTGTGTAAG36CAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCACGGGCACAGCGGTGAGAGCCTGAGGCTTCCCGG30CATCTCTGGAGCGCGTGACC 2340
Db 2281 GGCACGGGCACAGCGGTGAGAGCCTGAGGCTTCCCGG30CATCTCTGGAGCGCGTGACC 2340
QY 2341 GTGAGGTCTCTCCGTGGGAGATGACACGCGCCCGGTCGCGTACTCTTCTATCTG 2400
Db 2341 GTGAGGTCTCTCCGTGGGAGATGACACGCGCCCGGTCGCGTACTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTGTC 2460
Db 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTGTC 2460
QY 2461 TTGCACACAGCGTCTCAGGCTCTCCAAACAGGTGGAGCAGCGGACGACAGCCCCCT 2520
Db 2461 TTGCACACAGCGTCTCAGGCTCTCCAAACAGGTGGAGCAGCGGACGACAGCCCCCT 2520
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCCGTCGTCGCGGAGCTGGCGAGCTGGCGCTGCAG 2580
Db 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCCGTCGTCGCGGAGCTGGCGAGCTGGCGCTGCAG 2580
QY 2581 AAGCGGCGGTGGACTGTGCGGGCTCCGCGGCGAGCGGTCGCTGCTGTGATGCA 2640
Db 2581 AAGCGGCGGTGGACTGTGCGGGCTCCGCGGCGAGCGGTCGCTGCTGTGATGCA 2640
QY 2641 GCCATCGGCCGCTGGAGACACAGCTGCGGGAGCCCTGCCCCACCTGGAGCTCAGC 2700
Db 2641 GCCATCGGCCGCTGGAGACACAGCTGCGGGAGCCCTGCCCCACCTGGAGCTCAGC 2700
QY 2701 GCCTGGTCAACCTGCTCCAAAGAGCTGCGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGGTCAACCTGCTCCAAAGAGCTGCGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCGACGAGCGCGGCTGCTGGCGGAGCCAGTGCAACTTGCACCGCAAGCCCGCAG 2820
Db 2761 GTGGGCGACGAGCGCGGCTGCTGGCGGAGCCAGTGCAACTTGCACCGCAAGCCCGCAG 2820
QY 2821 GAGCTGACTTCTCGCTCCCTGAGGCGCTGTGA 2853
Db 2821 GAGCTGACTTCTCGCTCCCTGAGGCGCTGTGA 2853

RESULT 2

AAD35571

ID AAD35571 standard; cDNA; 3446 BP.

AC AAD35571;

XX AAD35571;

DT 26-JUL-2002 (first entry)

XX Human protease cDNA #4.

DE Human; novel human protein; NHP; protease; biological disorder; obesity;

KW high blood pressure; arthritis; connective tissue disorder; infertility;

KW gene therapy; enzyme; ss.

XX Homo sapiens.

OS Homo sapiens.

XX WO200226949-A2.

PN 04-APR-2002.

XX 27-SEP-2001; 2001WO-US30350.

XX 29-SEP-2000; 2000US-236689P.

PR (LEXI-) LEXICON GENETICS INC.

XX Friddle CJ, Hilbun E;

XX

PI

XX

DR

XX

PT

PT

XX

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity

Matches 2853;

Conservative

Mismatches

Indels

Gaps

0;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

WP1; 2002-372123/40.

Novel nucleic acid encoding a human protease, useful as a hybridization

probe for screening libraries and assessing gene expression patterns -

Disclosure; Page 40-41; 41pp; English.

The present sequence is a cDNA encoding novel human protein (NHP),

human protease. NHPs share structural similarity with animal proteases

particularly zinc metalloproteases. Sequences of the invention are

useful in therapeutic, diagnostic and pharmacogenomic applications.

NHP polynucleotides are used as hybridisation probes for screening

libraries and assessing gene expression patterns. They can also be

used for treating related biological disorders such as obesity, high

blood pressure, arthritis, connective tissue disorders and infertility.

They are also used in gene therapy.

Sequence 3446 BP; 612 A; 1114 C; 1101 G; 619 T; 0 other;

Query Match 100.0%; Score 2853; DB 24; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCCTTAACCCCTGGCTTTCCGCGGCGGAACCCCTGGAGGCTCTGAG 60

Db 397 ATGCTTCTGCTGGGATCCTTAACCCCTGGCTTTCCGCGGCGGAACCCCTGGAGGCTCTGAG 456

QY 61 CCAGAGCGGAGTAGTCTGCTTCCATCCGACTGGACCCGAGCATTAACGCGCGCGCTAC 120

Db 457 CCAGAGCGGAGTAGTCTGCTTCCATCCGACTGGACCCGAGCATTAACGCGCGCGCTAC 516

QY 121 TACTGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 180

Db 517 TACTGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 576

QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCCGCTTCC 240

Db 577 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGCTCCCGCTTCC 636

QY 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACTGGGAGCG 300

Db 637 ACTGAGCATCTGGGCTCCCGCTCCAGGGGCTCACCGGGGCTCTTTCAGACTGGGAGCG 696

QY 301 TGTCTTCTTCTGGGACGTGAACCGGAGCGGACTCGTTGCTGCTGAGCCTGTGC 360

Db 697 TGTCTTCTTCTGGGACGTGAACCGGAGCGGACTCGTTGCTGCTGAGCCTGTGC 756

QY 361 GGGGGCTCCGCGAGCCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC 420

Db 757 GGGGGCTCCGCGAGCCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC 816

QY 421 AATGCTAGGCGCGCGGCGGCGAGCGCAACAGCGGCGCACACTTCTCCAGCGCGCG 480

Db 817 AATGCTAGGCGCGCGGCGGCGAGCGCAACAGCGGCGCACACTTCTCCAGCGCGCG 876

QY 481 GGTGTTCCGGGCGGCTTCCGAGACCCACTCTCCCTCGGGGTGGCTTCCGGCTGG 540

Db 877 GGTGTTCCGGGCGGCTTCCGAGACCCACTCTCCCTCGGGGTGGCTTCCGGCTGG 936

QY 541 AACCCTCCCATCTCTACGGGCGCTTGAGACCTTACAAAGCGCGCGGCGGCTTCCGGGCG 600

Db 937 AACCCTCCCATCTCTACGGGCGCTTGAGACCTTACAAAGCGCGCGGCGGCTTCCGGGCG 996

QY 601 AGTCGTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660

Db 997 AGTCGTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1056

QY 661 GAGACGCTGGTGGTGGCGGAGTCAATGTGTCAGTTCCACGCGGCGGCGGCGGCGGCGG 720

Db 1057 GAGACGCTGGTGGTGGCGGAGTCAATGTGTCAGTTCCACGCGGCGGCGGCGGCGG 1116

QY 721 TATCTGCTGACGCTGCTGGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780

Db 1117 TATCTGTGACGCTGTGTGGCAAGGGGGGCGGACTCTACGGCCATCCCAAGACTCCTCAAC 1176
QY 781 CCATCAACATCTGTGGTCAAGGTGCTCTTCTTAGAGATGCTGACTCCGGGCCCAAG 840
Db 1177 CCCATCAACATCTGTGGTCAAGGTGCTCTTCTTAGAGATGCTGACTCCGGGCCCAAG 1236
QY 841 GTACCCGGCAATGGGCGCTGACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 900
Db 1237 GTACCCGGCAATGGGCGCTGACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTGAAC 1296
QY 901 AAAGTGAAGTGAAGACACCCCGAGTACTGGGACACTTGCCCATCTCTTCAACAGGCAAGGAC 960
Db 1297 AAAGTGAAGTGAAGACACCCCGAGTACTGGGACACTTGCCCATCTCTTCAACAGGCAAGGAC 1356
QY 961 CTGTGTGGGCCACCACTCTGTACACCTTGGGATGGCTGATGTGGTACCATGTCTGAC 1020
Db 1357 CTGTGTGGGCCACCACTCTGTACACCTTGGGATGGCTGATGTGGTACCATGTCTGAC 1416
QY 1021 CCCAAGAGAAGTCTGTCTGATTGAGGAGATGGGCTTCCATCAGCCTTCAACCACTGCC 1080
Db 1417 CCCAAGAGAAGTCTGTCTGATTGAGGAGATGGGCTTCCATCAGCCTTCAACCACTGCC 1476
QY 1081 CACGAGCTGGGCCACAGTGTTCACATGCCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1477 CACGAGCTGGGCCACAGTGTTCACATGCCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1536
QY 1141 TTTGGGAAGTCCGAGCAACCACTGATGTCGCCGACCTCATCCAGATCGACCGTGC 1200
Db 1537 TTTGGGAAGTCCGAGCAACCACTGATGTCGCCGACCTCATCCAGATCGACCGTGC 1596
QY 1201 AACCCCTGGTCAAGCTGCTGATGTCATCATCCAGACTTCTTGGACAGCGGGCAAGGT 1260
Db 1597 AACCCCTGGTCAAGCTGCTGATGTCATCATCCAGACTTCTTGGACAGCGGGCAAGGT 1656
QY 1261 GACTGCTCTGGACCAACCCAGCAAGCCATCTCTCTGCCGAGATCTGCCGGGGCC 1320
Db 1657 GACTGCTCTGGACCAACCCAGCAAGCCATCTCTCTGCCGAGATCTGCCGGGGCC 1716
QY 1321 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCT 1380
Db 1717 AGCTACACCTGAGCCAGCAGTGGAGCTGGCTTTTGGCTGGGCTCCAAAGCCCTGTCT 1776
QY 1381 TACATGCACTACTGACCAAGCTGTGTGCAACCGGAAGGCAAGGACAGATGTGTGC 1440
Db 1777 TACATGCACTACTGACCAAGCTGTGTGCAACCGGAAGGCAAGGACAGATGTGTGC 1836
QY 1441 CAGACCCGCCACTTCCCTGGGCGGATGGCACAGCTGTGGCGAGGGCAAGCTCTGCC 1500
Db 1837 CAGACCCGCCACTTCCCTGGGCGGATGGCACAGCTGTGGCGAGGGCAAGCTCTGCC 1896
QY 1501 AAAGGGCTCGTGGGAGACAACTCAACAAAGCACAGGCTGGATGTTCTTGGGCC 1560
Db 1897 AAAGGGCTCGTGGGAGACAACTCAACAAAGCACAGGCTGGATGTTCTTGGGCC 1956
QY 1561 AAATGGATCTTATGCCCCCTGCTCGGCACATGTGGTGGGCGTGACAGTGGCCAGG 1620
Db 1957 AAATGGATCTTATGCCCCCTGCTCGGCACATGTGGTGGGCGTGACAGTGGCCAGG 2016
QY 1621 AGGAGTGCACAAACCCACCCCTGCCAAGCGGGGCAAGTACTGCGAGGGAGTGAAGGTG 1680
Db 2017 AGGAGTGCACAAACCCACCCCTGCCAAGCGGGGCAAGTACTGCGAGGGAGTGAAGGTG 2076
QY 1681 AAATACCGATCTGATCTGGAGCCTTCCCAAGCTCAGCTCCGGAAGAGCTTCCGG 1740
Db 2077 AAATACCGATCTGATCTGGAGCCTTCCCAAGCTCAGCTCCGGAAGAGCTTCCGG 2136
QY 1741 GAGGAGCAGTGTAGGCTTTTCAACGGGTACAAACACAGCACCAACCGCTCACTCTGCC 1800
Db 2137 GAGGAGCAGTGTAGGCTTTTCAACGGGTACAAACACAGCACCAACCGCTCACTCTGCC 2196
QY 1801 GTGGCATGGTGGCCCAAGTACTCGGGCTGTCTCCCGGGGACAAAGTGCAGCTCATCTGC 1860
|||||

Db 2197 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGACAACTGCAAGCTCACTGC 2256
QY 1861 CGAGCAATGGCACTGGCTACTTCTATGTCTGGCACCCCAAGTGGTGGAGCGCACCTG 1920
Db 2257 CGAGCAATGGCACTGGCTACTTCTATGTCTGGCACCCCAAGTGGTGGAGCGCACCTG 2316
QY 1921 TGCTCTCTGACTCCACCTCGTCTGTCCAAGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 2317 TGCTCTCTGACTCCACCTCGTCTGTCCAAGCAAGTGCATCAAGGCTGGCTGTGAT 2376
QY 1981 GGGACCTGGGCTCCAAGAAGAGATTGCACAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
Db 2377 GGGACCTGGGCTCCAAGAAGAGATTGCACAAGTGTGGGGTGTGGGGGAGACAATAAG 2436
QY 2041 AGCTCAAGAAGGTGACTGGACTCTTCCACCAAGCCCATGCTGCTCAATTTCTGGTG 2100
Db 2437 AGCTCAAGAAGGTGACTGGACTCTTCCACCAAGCCCATGCTGCTCAATTTCTGGTG 2496
QY 2101 GCCATCCCGGAGGGCGCTCAAGCATCGACATCGCCAGCGCGGTACAAAGGGCTGATC 2160
Db 2497 GCCATCCCGGAGGGCGCTCAAGCATCGACATCGCCAGCGCGGTACAAAGGGCTGATC 2556
QY 2161 GGGGATGACAATCTACTGGCTCTGAAGAAACAGCAAGCAAGCAAGTACCTGCTCAAGGGCAT 2220
Db 2557 GGGGATGACAATCTACTGGCTCTGAAGAAACAGCAAGCAAGCAAGTACCTGCTCAAGGGCAT 2616
QY 2221 TTCTGGTGTGGCGGTGGAGCGGACCTGTGTGAAGGGACGTCTCTCGGGTATAGC 2280
Db 2617 TTCTGGTGTGGCGGTGGAGCGGACCTGTGTGAAGGGACGTCTCTCGGGTATAGC 2676
QY 2281 GGCAGGGCACAGCGGTGGAGAGCTGCAGGCTTCCCGGCCCATCTCGAGCGCTGACC 2340
Db 2677 GGCAGGGCACAGCGGTGGAGAGCTGCAGGCTTCCCGGCCCATCTCGAGCGCTGACC 2736
QY 2341 GTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2737 GTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2796
QY 2401 CCCAAGAGCCTCGGAGAGCAAGTCTCTCATCCCAAGAGCCCGCGGGGACCTCTGTC 2460
Db 2797 CCCAAGAGCCTCGGAGAGCAAGTCTCTCATCCCAAGAGCCCGCGGGGACCTCTGTC 2856
QY 2461 TTGCACAACAGCGTCTCAGCCTCTCCAAACAGTGGAGCAGCCGACGAGCGCCCT 2520
Db 2857 TTGCACAACAGCGTCTCAGCCTCTCCAAACAGTGGAGCAGCCGACGAGCGCCCT 2916
QY 2521 GCACGCTGGGTGGTGGCAGCTGGGGCGCGTCTCCGGAGCTGGCGGAGTGGCTGACG 2580
Db 2917 GCACGCTGGGTGGTGGCAGCTGGGGCGCGTCTCCGGAGCTGGCGGAGTGGCTGACG 2976
QY 2581 AAGCGGCGGTGACTGTGGGGCTCCCGGGCAGCGCAGCTCCCTGCCTGTGAAGCA 2640
Db 2977 AAGCGGCGGTGACTGTGGGGCTCCCGGGCAGCGCAGCGTCCCTGCCTGTGAAGCA 3036
QY 2641 GGCATCGGCCCGTGGAGACAAAGCTTCCGGGAGCCCTGCCACCTGGGAGCTGACG 2700
Db 3037 GGCATCGGCCCGTGGAGACAAAGCTTCCGGGAGCCCTGCCACCTGGGAGCTGACG 3096
QY 2701 GCCTGTGCTACCTTGTCTCAAGAGCTGGCGCGGGGATTTCAAGAGGCTCACTCAAGTGT 2760
Db 3097 GCCTGTGCTACCTTGTCTCAAGAGCTGGCGCGGGGATTTCAAGAGGCTCACTCAAGTGT 3156
QY 2761 GTGGGCAAGGAGCGGCTGTGTGCCCGGGACCAAGTGCACCTGCACCGCAAGCCCCAG 2820
Db 3157 GTGGGCAAGGAGCGGCTGTGTGCCCGGGACCAAGTGCACCTGCACCGCAAGCCCCAG 3216
QY 2821 GAGCTGAGCTTCTGGCTCTGAGCGCTGCTGA 2853
|||||
Db 3217 GAGCTGAGCTTCTGGCTCTGAGCGCTGCTGA 3249
|||||

RESULT 3
AAH41003

Db	1377	TCCTTACATGCACTACTGACCAAGCTGTGGTGCACCGGTAAGGCCAAGGACAGATGCT	1436
Qy	1437	GTGCCAGACCCGCACTTCCCTTGGCCGATGGCACCAGCTGTGGCGAGGGCAAGCTCTG	1496
Db	1437	GTGCCAGACCCGCACTTCCCTTGGCCGATGGCACCAGCTGTGGCGAGGGCAAGCTCTG	1496
Qy	1497	CCTCAAGGGGCTGGTGGAGAGACACAACTTCAACAGCACAGGCTGATGGTTCCTG	1556
Db	1497	CCTCAAGGGGCTGGTGGAGAGACACAACTTCAACAGCACAGGCTGATGGTTCCTG	1556
Qy	1557	GGCCAAATGGATCCCTATGGCCCTGCTGCGGCACATGTGTGGGGGCTGCAGCTGGC	1616
Db	1557	GGCCAAATGGATCCCTATGGCCCTGCTGCGGCACATGTGTGGGGGCTGCAGCTGGC	1616
Qy	1617	CAGGAGCAGTGCACCAACCCCTCCCTGCAACAGGGGTAAGTACTGCGAGGAGTGA	1676
Db	1617	CAGGAGCAGTGCACCAACCCCTCCCTGCAACAGGGGTAAGTACTGCGAGGAGTGA	1676
Qy	1677	GGTGAATACCGATCCCTGCAATCTGGAGCCCTGCCCTGAGCTGAGCTTCCGGAAGAGCTT	1736
Db	1677	GGTGAATACCGATCCCTGCAATCTGGAGCCCTGCCCTGAGCTGAGCTTCCGGAAGAGCTT	1736
Qy	1737	CCGGGAGGAGTGTGAGGCTTTCAACGCTTACACGCTTACACGAGTCCGGAAGAGCTT	1796
Db	1737	CCGGGAGGAGTGTGAGGCTTTCAACGCTTACACGCTTACACGAGTCCGGAAGAGCTT	1796
Qy	1797	CGCGTGCGATGGTCCCAAGTACTCCGGCTGTCTCCCGGACAAAGTGCAGCTCAT	1856
Db	1797	CGCGTGCGATGGTCCCAAGTACTCCGGCTGTCTCCCGGACAAAGTGCAGCTCAT	1856
Qy	1857	CTGCCGAGCAATGGCACTGGCTACTTCTATGTCTGGCACCCCAAGTGTGACGGCAC	1916
Db	1857	CTGCCGAGCAATGGCACTGGCTACTTCTATGTCTGGCACCCCAAGTGTGACGGCAC	1916
Qy	1917	GCTGTGCTCTCTGACTCCACTCCCTGCTGTCTCCAGGCAAGTGCATCAGCTGGCTG	1976
Db	1917	GCTGTGCTCTCTGACTCCACTCCCTGCTGTCTCCAGGCAAGTGCATCAGCTGGCTG	1976
Qy	1977	TGATGGGAACCTGGCTCCAAAGAGAGATTCGACAACTGGGCTGTGTGGGGGAGACAA	2036
Db	1977	TGATGGGAACCTGGCTCCAAAGAGAGATTCGACAACTGGGCTGTGTGGGGGAGACAA	2036
Qy	2037	TAAGAGCTGCAAGAGTGTGACTGCTCTTCAACCAAGCCATGATGGCTACAAATTCGT	2096
Db	2037	TAAGAGCTGCAAGAGTGTGACTGCTCTTCAACCAAGCCATGATGGCTACAAATTCGT	2096
Qy	2097	GGTGCCATCCCGAGCGCTCAAGCATCGACATCCGCGAGCGGTTACAAAGGCT	2156
Db	2097	GGTGCCATCCCGAGCGCTCAAGCATCGACATCCGCGAGCGGTTACAAAGGCT	2156
Qy	2157	GATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGGCAAGTACCTGTCAACGG	2216
Db	2157	GATCGGGATGACAACTACCTGGCTCTGAAGAACAGCAAGGCAAGTACCTGTCAACGG	2216
Qy	2217	GCATTTCTGTGTGGGGTGGAGCGGACCTGGTGAAGGCAAGTGTGCTGCGGTA	2276
Db	2217	GCATTTCTGTGTGGGGTGGAGCGGACCTGGTGAAGGCAAGTGTGCTGCGGTA	2276
Qy	2277	CAGCGGCAGCGGACAGCGGTGGAGAGCTTCCAGGCTTCCCGGCCATCTGTGAGCGGCT	2336
Db	2277	CAGCGGCAGCGGACAGCGGTGGAGAGCTTCCAGGCTTCCCGGCCATCTGTGAGCGGCT	2336
Qy	2337	GACCTGGAGTGTCTCTCGGTGGGGAAGATGACACCGCGCGGTGCTCTCTTCTA	2396
Db	2337	GACCTGGAGTGTCTCTCGGTGGGGAAGATGACACCGCGCGGTGCTCTCTTCTA	2396
Qy	2397	TCTGCCCAAGAGCTCCGGAGGACAAGTCTCTATCCCAAGGACCCCGGGACCCCTC	2456
Db	2397	TCTGCCCAAGAGCTCCGGAGGACAAGTCTCTATCCCAAGGACCCCGGGACCCCTC	2456
Qy	2457	TGCTTTGACACACAGCTTCTCAGCTCTTCCAAACAGGTGGAGCGCGGACGAGGCC	2516
Db	2457	TGCTTTGACACACAGCTTCTCAGCTCTTCCAAACAGGTGGAGCGCGGACGAGGCC	2516

Qy	2517	CCCTGACGCTGGTGGCTGGCAGCTGGGGCCGCTGCTCCGCGAGCTGCGGCAAGTGGCCT	2576
Db	2517	CCCTGACGCTGGTGGCTGGCAGCTGGGGCCGCTGCTCCGCGAGCTGCGGCAAGTGGCCT	2576
Qy	2577	GCAGAAGCGGCGGTGGACTGCGGGGCTCGCGGGGAGCGGCTGCGCTGTGA	2636
Db	2577	GCAGAAGCGGCGGTGGACTGCGGGGCTCGCGGGGAGCGGCTGCGCTGTGA	2636
Qy	2637	TGCAGCCATCGCCCGTGGAGACACAGCTGCGGGGAGCCCTGCCACCTGGAGCT	2696
Db	2637	TGCAGCCATCGCCCGTGGAGACACAGCTGCGGGGAGCCCTGCCACCTGGAGCT	2696
Qy	2697	CAGCGCTTGGTCAACCTGCTCCAGAGCTGCGGCGGGGATTCAGAGGCGCTCACTCAA	2756
Db	2697	CAGCGCTTGGTCAACCTGCTCCAGAGCTGCGGCGGGGATTCAGAGGCGCTCACTCAA	2756
Qy	2757	GTGTGTGGGCGACGAGCGCGCTGCTGGCGGGGAGCGGCTGCACTGCAAGGCC	2816
Db	2757	GTGTGTGGGCGACGAGCGCGCTGCTGGCGGGGAGCGGCTGCACTGCAAGGCC	2816
Qy	2817	CCAGGAGCTGGACTTCTGCTGCTGAGGCGCTGTGA	2853
Db	2817	CCAGGAGCTGGACTTCTGCTGCTGAGGCGCTGTGA	2853

RESULT 4
ABK12894
ID ABK12894 standard; cDNA; 2930 BP.
XX
AC ABK12894;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-11 cDNA sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT /*tag= a
FT /partial
FT /product= "Human protease PRTS-11"
FT /note= "This sequence lacks a stop codon"

WO200198468-A2.
27-DEC-2001.
13-JUN-2001; 2001WO-US19178.
16-JUN-2000; 2000US-212336P.
22-JUN-2000; 2000US-213955P.
29-JUN-2000; 2000US-215396P.
07-JUL-2000; 2000US-216821P.
14-JUL-2000; 2000US-218946P.
(INCY-) INCYTE GENOMICS INC.
Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;
Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
Wallia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
Azimzal Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
Kaillick DA;
WPI; 2002-090437/12.

with sequence

DR P-PSDB; AAU74751.
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell
PT proliferative (e.g. cancer) disorders -
XX Claim 5; Page 168; 177pp; English.
XX The present invention relates to twenty one new human proteases,
CC referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and
CC polypeptides of the invention are useful in the diagnosis, treatment and
CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
CC myocardial infarction, autoimmune/inflammatory e.g. acquired
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and
CC endometriosis disorders. Numerous other examples of each disorder are
CC given in the specification. The present nucleic acid sequence encodes
XX the human protease PRTS-11 protein of the invention.
SQ Sequence 2930 BP; 529 A; 958 C; 931 G; 512 T; 0 other;
Query Match 77.0%; Score 2196; DB 24; Length 2930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 ATGCTTCTGCTGGGATCTAACCCTGGCTTTCCGCGGCGAAGCGCTGAGGCTCTGAG 60
DB 75 ATGCTTCTGCTGGGATCTAACCCTGGCTTTCCGCGGCGAAGCGCTGAGGCTCTGAG 134
QY 61 CGAGCGGGAGGTAGTCTTCCATCCGACTGGACCGCGACATTAAGCGCGCGGCTAC 120
DB 135 CGAGCGGGAGGTAGTCTTCCATCCGACTGGACCGCGACATTAAGCGCGCGGCTAC 194
QY 121 TACTGCGGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 180
DB 195 TACTGCGGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCACAGCATTT 254
QY 181 CAGGAGGACTTTTACCTACCTGACCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 255 CAGGAGGACTTTTACCTACCTGACCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 314
QY 241 ACTGAGCATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 315 ACTGAGCATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374
QY 301 TGCTTCTATTTGCGGAGCTGAACCGCGAGCGGACTCGTTCGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 375 TGCTTCTATTTGCGGAGCTGAACCGCGAGCGGACTCGTTCGCTGCTGCTGCTGCTGCTGCTGCT 434
QY 361 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGCGCGGAGTGTGCTATGCTATGCTGCTGCTGCTGCTGCT 420
DB 435 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGCGCGGAGTGTGCTATGCTATGCTATGCTATGCTATGCT 494
QY 421 AATGCTAGCGCGCGCGGCGGAGCGACACAGCGAGCGGCGACACTTCTCCAGCGCGG 480
DB 495 AATGCTAGCGCGCGCGGCGGAGCGACACAGCGAGCGGCGACACTTCTCCAGCGCGG 554
QY 481 GGTGTTCCGGGGGGGCTTTCCGAGAGACCCACTCTCGCTGCGGGTGGCTCGGGCTGG 540
DB 555 GGTGTTCCGGGGGGGCTTTCCGAGAGACCCACTCTCGCTGCGGGTGGCTCGGGCTGG 614
QY 541 AACCCCGCCATCTACGGGCCCTGGACCTTACAGCGCGCGGCGGCGGCTTCGGGAG 600
DB 615 AACCCCGCCATCTACGGGCCCTGGACCTTACAGCGCGCGGCGGCGGCTTCGGGAG 674
QY 601 AGTCGTAGCGCGCGAGTCTGGGCGCGCAGCGCTTTCGTCTATCTCCCGCGGTACGTG 660
DB 675 AGTCGTAGCGCGCGAGTCTGGGCGCGCAGCGCTTTCGTCTATCTCCCGCGGTACGTG 734

Db	1815	GAGGACAGTGTGAGGCTTCAACGGCTACAAACACAGCACCAGCGCTCACTCTCGCC	1874
Qy	1801	GTGCATGGGTCACAGTACTTCGGCGGTGTCTCCCGGGACAGTGCACAGCTCATCTGC	1860
Db	1875	GTGCATGGGTCACAGTACTTCGGCGGTGTCTCCCGGGACAGTGCACAGCTCATCTGC	1934
Qy	1861	CGACCAATGACGCTGCTCTATGTGTGTGACCCAA--GGTGTGGACGGCAGC	1917
Db	1935	CGACCAATGACGCTGCTCTATGTGTGTGACCCAA--GGTGTGGACGGCAGC	1994
Qy	1918	CTGTCTCTCTGACTCCACCTCGCTGTGTCCAAAGCAAGTGCATCAAGGTGGCTGT	1977
Db	1995	CTGTCTCTCTGACTCCACCTCGCTGTGTCCAAAGCAAGTGCATCAAGGTGGCTGT	2054
Qy	1978	GATGGACACTGGGCTCCAAAGAGATTCGACAGTGTGGGTGTGTGGGGAGACAAT	2037
Db	2055	GATGGACACTGGGCTCCAAAGAGATTCGACAGTGTGGGTGTGTGGGGAGACAAT	2114
Qy	2038	AAGAGCTCAAGAAGGTGACTGAGTCTCTTCCAAAGCCCATGCTATGCTACAATTTCGTG	2097
Db	2115	AAGAGCTCAAGAAGGTGACTGAGTCTCTTCCAAAGCCCATGCTATGCTACAATTTCGTG	2174
Qy	2098	GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCGCCAGCGCGGTTCACAAGGCGTG	2157
Db	2175	GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCGCCAGCGCGGTTCACAAGGCGTG	2234
Qy	2158	ATCGGGGATGACACTACTCTGGCTCTGAAGAACAGCAAGTACTGCTCAACGGG	2217
Db	2235	ATCGGGGATGACACTACTCTGGCTCTGAAGAACAGCAAGTACTGCTCAACGGG	2294
Qy	2218	CATTTCTGGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCACTGTCTCGGTAC	2277
Db	2295	CATTTCTGGTGTGGCGGTGGAGCGGACCTGTGTGTGAAGGCACTGTCTCGGTAC	2354
Qy	2278	ACGGCAGCGACAGCGTGGAGGCTGACAGCTTCCCGGCCCATCTCTGAGCGCGTG	2337
Db	2355	ACGGCAGCGGACAGCGGTGGAGGCTGACAGCTTCCCGGCCCATCTCTGAGCGCGTG	2414
Qy	2338	ACCGTGGAGGCTCTCTCGTGGGGAAGATGACACCGCCCGGTCGCTCTCTAT	2397
Db	2415	ACCGTGGAGGCTCTCTCGTGGGGAAGATGACACCGCCCGGTCGCTCTCTAT	2474
Qy	2398	CTGCCCAAGAGCCTCGGAGGACAACTCCTCATCCC	2436
Db	2475	CTGCCCAAGAGCCTCGGAGGACAACTCCTCATCCC	2513
RESULT 5			
AAS97182			
Id	AAS97182 standard; cDNA; 2937 BP.		
XX	AAS97182;		
AC	AAS97182;		
XX	26-FEB-2002 (first entry)		
DT	Human metalloprotease partial DNA sequence #11.		
DE	Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;		
XX	vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;		
KW	hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;		
KW	anorectic; antinflammatory; aspartyl protease; cysteine protease;		
KW	metalloprotease; serine protease; cancer; haematopoietic; breast; colon;		
KW	lung; prostrate; cervical; brain; ovarian; bladder; kidney; pain;		
KW	immune-related disease; cardiovascular disease; neuronal disease;		
KW	migraine; sexual dysfunction; mood disorder; attention disorder;		
KW	cognition disorder; hypotension; hypertension; psychotic disorder;		
KW	dyskinesia; metabolic disorder; inflammatory disorder; ss.		
OS	Homo sapiens.		
XX	W0200183782-A2.		
PN			
XX			

PD	08-NOV-2001.	
XX	04-MAY-2001; 2001WO-US14431.	
XX	04-MAY-2000; 2000US-201879P.	
XX	(SUGE-) SUGEN INC.	
PA	Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;	
PI	Payne V;	
XX	WPI; 2002-041502/05.	
DR	P-PSDB; AA072899.	
XX	Novel protease polypeptide useful for screening for substances that may	
PT	be used to treat, e.g., cancers, immune-related diseases,	
PT	cardiovascular disease, migraine, pain, psychotic and inflammatory	
PT	disorders -	
XX	Claim 30; Figure 1P; 232pp; English.	
PS	The invention relates to an isolated, enriched, or purified protease	
CC	polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to	
CC	screen for substances (S) that may modulate its activity. Administering	
CC	S (which modulates protease activity in vitro) may be used to treat a	
CC	disease or disorder selected from cancers (e.g., of tissues, of blood or	
CC	haematopoietic origin, of the breast, colon, lung, prostate, cervical,	
CC	brain, ovarian, bladder or kidney), immune-related diseases and	
CC	disorders, cardiovascular disease, brain or neuronal-associated diseases	
CC	(e.g., central or peripheral nervous system diseases, migraine, pain,	
CC	sexual dysfunction, mood disorders, attention disorders, cognition	
CC	disorders, hypotension, hypertension, psychotic disorders, neurological	
CC	disorders and dyskinesias), metabolic disorders and inflammatory	
CC	disorders. (I) may also be useful as a diagnostic tool for a disease or	
CC	disorder such as those above. AAS97159-AAS97195 represent human	
CC	protease coding sequences and primers of the invention.	
XX	Sequence 2937 BP; 531 A; 958 C; 936 G; 512 T; 0 other;	
SQ	Query Match 77.0%; Score 2196; DB 24; Length 2937;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;	
QY	1 ATGCTTCTGCTGGCATCTCAACCTTGGCTTTTCGGCGGGCAACCGCTCGAGGCTCTAG	60
DB	76 ATGCTTCTGCTGGCATCTCAACCTTGGCTTTTCGGCGGGCAACCGCTCGAGGCTCTAG	135
QY	61 CCAGAGCGGAGGTAGTGTTCCTCCATCCGACTCGGACATTAACGCCGCCCTAC	120
DB	136 CCAGAGCGGAGGTAGTGTTCCTCCATCCGACTCGGACATTAACGCCGCCCTAC	195
QY	121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCAGCATTT	180
DB	196 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTTCAGATCAGCATTT	255
QY	181 CAGGAGGACTTTTACCTTACACCTGACCGCGATGCTCAGTTCTTGGCTCCCGCTTCTCC	240
DB	256 CAGGAGGACTTTTACCTTACACCTGACCGCGATGCTCAGTTCTTGGCTCCCGCTTCTCC	315
QY	241 ACTGAGCATCTGGGCTGCCCTCCAGGGCTCACCGGGGCTCTTCAGACCTGCGACCG	300
DB	316 ACTGAGCATCTGGGCTGCCCTCCAGGGCTCACCGGGGCTCTTCAGACCTGCGACCG	375
QY	301 TGCTTCTATTCTGGGACGTGAACCGGAGCTCGTTTCGCTGCTGTGAGCCTGTGC	360
DB	376 TGCTTCTATTCTGGGACGTGAACCGGAGCTCGTTTCGCTGCTGTGAGCCTGTGC	435
QY	361 GGGGGCTCCCGAGGCTTTTGGCTACCGAGCGCGAGTATGTATTCACCGCGTGC	420
DB	436 GGGGGCTCCCGAGGCTTTTGGCTACCGAGCGCGAGTATGTATTCACCGCGTGC	495
QY	421 AATGCTAGCGCGCGGCGGCGACCGCAACAGCCAGGCGCACACCTTCTCCAGCGCGG	480

Db 496 AATGCTAGCGCGCGCGCGAGCGCAACAGCCAGGCGGCACACCTTCTCCAGCGCGG 555
QY 481 GTGTTCGGGGGGGCTTCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGGTGG 540
Db 556 GTGTTCGGGGGGGCTTCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGGTGG 615
QY 541 AACCCGGCATTCTTACGGGCCCTTGAGCCCTTACAGCGCGCGCGGGGCTTCGGGGAG 600
Db 616 AACCCGGCATTCTTACGGGCCCTTGAGCCCTTACAGCGCGCGCGGGGCTTCGGGGAG 675
QY 601 AGTCGTAGCGCGCGAGCTGTGGGGCGCCAAAGGTTTGTGTCTATCCCGGTAGCTG 660
Db 676 AGTCGTAGCGCGCGAGCTGTGGGGCGCCAAAGGTTTGTGTCTATCCCGGTAGCTG 735
QY 661 GAGAGCTGGTGGTCCGCGAGAGTCAATAGTTCAGTTCACCGCGCGGAGCTTGGAAAT 720
Db 736 GAGAGCTGGTGGTCCGCGAGAGTCAATAGTTCAGTTCACCGCGCGGAGCTTGGAAAT 795
QY 721 TATCTGCTACGCTGTGCGACGCGCGGCGACTCTACCGCCATCCCGATCCTCAAC 780
Db 796 TATCTGCTACGCTGTGCGACGCGCGGCGACTCTACCGCCATCCCGATCCTCAAC 855
QY 781 CCATCAACATGTTGTGGTCAAGTGTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 840
Db 856 CCATCAACATGTTGTGGTCAAGTGTGCTTCTTAGAGATCGTACTCCGGGCCCAAG 915
QY 841 GTCCGGGCAANTGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAACTGAAC 900
Db 916 GTCCGGGCAANTGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAACTGAAC 975
QY 901 AAAGTGAAGTGAACAGCACCCCGAGTACTGGGACACTGCGCATCTTTCACCGAGCAGAC 960
Db 976 AAAGTGAAGTGAACAGCACCCCGAGTACTGGGACACTGCGCATCTTTCACCGAGCAGAC 1035
QY 961 CTGTGTGGAGCACACACTGTGACACCTGGGCATGGCTATGTGGGTACCATGTGTGAC 1020
Db 1036 CTGTGTGGAGCACACACTGTGACACCTGGGCATGGCTATGTGGGTACCATGTGTGAC 1095
QY 1021 CCCAAGAGAAGTGTCTGTCAATTGAGGACGATGGGCTTCCATCAGCCTTCACCACTGCC 1080
Db 1096 CCCAAGAGAAGTGTCTGTCAATTGAGGACGATGGGCTTCCATCAGCCTTCACCACTGCC 1155
QY 1081 CAGGAGCTGGGCGACGTGTTCAACATGCCCCATGACATGTGAAGTGTGTGAGGAGGTG 1140
Db 1156 CAGGAGCTGGGCGACGTGTTCAACATGCCCCATGACATGTGAAGTGTGTGAGGAGGTG 1215
QY 1141 TTTGGGAAGCTCGGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGC 1200
Db 1216 TTTGGGAAGCTCGGAGCAACACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGC 1275
QY 1201 AACCCCTGGTACGCTGCGAGTGTGCGATCATCACCGACTTCTGGAGACGGGCGACGGT 1260
Db 1276 AACCCCTGGTACGCTGCGAGTGTGCGATCATCACCGACTTCTGGAGACGGGCGACGGT 1335
QY 1261 GACTGCTCTCTGACCAACCCAGCAAGCCATCTCCCTGCCGAGGATCTGCCGGGCGCC 1320
Db 1336 GACTGCTCTCTGACCAACCCAGCAAGCCATCTCCCTGCCGAGGATCTGCCGGGCGCC 1395
QY 1321 AGCTACACCCCTGAGCAGCAGTGGAGTGGCTTTTGGGCTGGGCTTCCAAACCCCTGTGCT 1380
Db 1396 AGCTACACCCCTGAGCAGCAGTGGAGTGGCTTTTGGGCTGGGCTTCCAAACCCCTGTGCT 1455
QY 1381 TACATCAGTACTGCAACCAAGTGTGTGCACCGGGAAGGCCAAGGACACATGTGTGC 1440
Db 1456 TACATCAGTACTGCAACCAAGTGTGTGCACCGGGAAGGCCAAGGACACATGTGTGC 1515
QY 1441 CAGACCCGCACTTCCCTGGGCGGATGGCACCCAGCTGTGGCGAGGCAAGCTCTGGCTC 1500
Db 1516 CAGACCCGCACTTCCCTGGGCGGATGGCACCCAGCTGTGGCGAGGCAAGCTCTGGCTC 1575
QY 1501 AAAGGGGCTCTGCTGGAGAGACACACCTCAACAGACGACAGGCTGGATGTTCTTGGGCC 1560
Db 1576 AAAGGGGCTCTGCTGGAGAGACACACCTCAACAGACGACAGGCTGGATGTTCTTGGGCC 1635

QY 1561 AAATGGATCCCTATGTCCTGTGTCGCGACATGTGTGGGGCGTGCAGCTGGCCAG 1620
Db 1636 AAATGGATCCCTATGTCCTGTGTCGCGACATGTGTGGGGCGTGCAGCTGGCCAG 1695
QY 1621 AGGCAAGTGCACCAACCCCTGCGCAACCGGGGCAAGTACTCGAGGAGTGAAGGTG 1680
Db 1696 AGGCAAGTGCACCAACCCCTGCGCAACCGGGGCAAGTACTCGAGGAGTGAAGGTG 1755
QY 1681 AAATACCGATCCTGCAATCTGAGCCCTGCCCGCAGCTCAGCCTCCGGAAGAGCTTCCCG 1740
Db 1756 AAATACCGATCCTGCAATCTGAGCCCTGCCCGCAGCTCAGCCTCCGGAAGAGCTTCCCG 1815
QY 1741 GAGGAGAGTGTGAGGCTTTCAACGGCTTAAACACAGCAGCAACCGGCTCACTCTGCC 1800
Db 1816 GAGGAGAGTGTGAGGCTTTCAACGGCTTAAACACAGCAGCAACCGGCTCACTCTGCC 1875
QY 1801 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCGGGCAAGTGCAGCTCATCTGC 1860
Db 1876 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCGGGCAAGTGCAGCTCATCTGC 1935
QY 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGACCCAA ---GGTGTGGAGCGGACG 1917
Db 1936 CGAGCAATGGCACTGGCTTCTATGTGTGGACCCAAAGTGTGTGGAGCGGACG 1995
QY 1918 CTGTGCTCTCCTGACTCCACCTCGGTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGT 1977
Db 1996 CTGTGCTCTCCTGACTCCACCTCGGTGTGTCCAAAGCAAGTGCATCAAGGCTGGCTGT 2055
QY 1978 GATGGAACCTGGGCTCCAAAGAGATTTCGACAGTGTGGGTGTGTGGGGGAGACAAT 2037
Db 2056 GATGGAACCTGGGCTCCAAAGAGATTTCGACAGTGTGGGTGTGTGGGGGAGACAAT 2115
QY 2038 AAGAGCTCAAGAGTGAAGTGTGACTTTCACAGCCCATGATGCTACAATTTCTGT 2097
Db 2116 AAGAGCTCAAGAGTGAAGTGTGACTTTCACAGCCCATGATGCTACAATTTCTGT 2175
QY 2098 GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCCAGCGGTTACAAAGGCTG 2157
Db 2176 GTGGCCATCCCGCAGGCGCTCAAGCATCGACATCCGCCAGCGGTTACAAAGGCTG 2235
QY 2158 ATCGGGATGACAACTACTGCTCTGAAGACAGCAAGCAAGTACTCTCAAGCGG 2217
Db 2236 ATCGGGATGACAACTACTGCTCTGAAGACAGCAAGCAAGTACTCTCAAGCGG 2295
QY 2218 CATTTCTGTGTGTCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTCGGTAC 2277
Db 2296 CATTTCTGTGTGTCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTCGGTAC 2355
QY 2278 AGCGGCACGCGCACAGCGGTGGAGCGCTTCAGAGGCTTCCCGGCCCATGCTGAGCGCGT 2337
Db 2356 AGCGGCACGCGCACAGCGGTGGAGCGCTTCAGAGGCTTCCCGGCCCATGCTGAGCGCGT 2415
QY 2338 ACCTGGAGTCTCTCTCGGTGGGGAAGTGAACACCGCCCGGGTCCGCTACTCTTCTAT 2397
Db 2416 ACCTGGAGTCTCTCTCGGTGGGGAAGTGAACACCGCCCGGGTCCGCTACTCTCTAT 2475
QY 2398 CTGCCCCAAGAGCTCTGGGAGGACAAAGTCTCTATCCC 2436
Db 2476 CTGCCCCAAGAGCTCTGGGAGGACAAAGTCTCTATCCC 2514

RESULT 6

AAD35570

ID AAD35570 standard; cDNA; 1104 BP.

XX AAD35570;

XX AC

XX DT 26-JUL-2002 (first entry)

XX DE Human protease cDNA #3.

XX XX

KW Human; novel human protein; NHP; protease; biological disorder; obesity;

high blood pressure; arthritis; connective tissue disorder; infertility;
gene therapy; enzyme; gene; ss.
Homo sapiens.

Key Location/Qualifiers
CDS 1..1104
/*tag= a
/product= "Human protease #3"

WO200226949-A2.

04-APR-2002.

27-SEP-2001; 2001WO-US30350.

29-SEP-2000; 2000US-236689P.

(LEXI-) LEXICON GENETICS INC.

Friddle CJ, Hilbun E;

WPI; 2002-372123/40.

P-PSDB; AAE22542.

Novel nucleic acid encoding a human protease, useful as a hybridization probe for screening libraries and assessing gene expression patterns -
Disclosure; Page 38-39; 41pp; English.

The present sequence is a cDNA encoding novel human protein (NHP), human protease. NHPs share structural similarity with animal proteases particularly zinc metalloproteases. Sequences of the invention are useful in therapeutic, diagnostic and pharmacogenomic applications. NHP polynucleotides are used as hybridisation probes for screening libraries and assessing gene expression patterns. They can also be used for treating related biological disorders such as obesity, high blood pressure, arthritis, connective tissue disorders and infertility. They are also used in gene therapy.

Sequence 1104 BP; 189 A; 357 C; 345 G; 213 T; 0 other;

Query Match 38.2%; Score 1091; DB 24; Length 1104;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGCATCTAAACCTTGGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGCATCTAAACCTTGGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGAGCCGACATTAACGGCCCGGCTAC 120
DB 61 CCAGAGCGGGAGGTAGTGTCTCCATCCGACTGGAGCCGACATTAACGGCCCGGCTAC 120
QY 121 TACTGGCGGGTCCGGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCAGACATTT 180
DB 121 TACTGGCGGGTCCGGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCAGACATTT 180
QY 181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTTTGGCTCCCGCTTCTCC 240
DB 181 CAGGAGGACTTTTACCTACACTGACCGCGGATGCTCAGTTCTTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGGTCCCGCTCCAGGGGTCAACGGGGGCTCTTTCAGACCTGCGACGC 300
DB 241 ACTGAGCATCTGGGGTCCCGCTCCAGGGGTCAACGGGGGCTCTTTCAGACCTGCGACGC 300
QY 301 TGCCTTCTATCTGGGAGCTGAACCGCGGAGCTCGTTTCGCTGTGTGAGCTGTGC 360
DB 301 TGCCTTCTATCTGGGAGCTGAACCGCGGAGCTCGTTTCGCTGTGTGAGCTGTGC 360
QY 361 GGGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCGGAGTATGTCATTTAGCCCGCTGCC 420
DB 361 GGGGGGCTCCCGGAGGCTTTGGCTACCGAGCGCGGAGTATGTCATTTAGCCCGCTGCC 420

QY 421 AATGCTAGCGCGCGGGCGCAGCGCAACAGCGGCGCAGACCTTCTCCAGCGCGG 480
DB 421 AATGCTAGCGCGCGGGCGCAGCGCAACAGCGGCGCAGACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCCGGCGGGGCTTCCGGAGAGCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
DB 481 GGTGTTCCGGCGGGGCTTCCGGAGAGCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
QY 541 AACCCCGCATCTACGGGCGCTTACAGCCCTTACAGCGCGGCGGGGCTTCCGGGAG 600
DB 541 AACCCCGCATCTACGGGCGCTTACAGCCCTTACAGCGCGGCGGGGCTTCCGGGAG 600
QY 601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCCAAAGCGTTTGTGTCTATCCCGGTAGCTG 660
DB 601 AGTCGTAGCGCGCGAGGTCTGGGCGCGCCAAAGCGTTTGTGTCTATCCCGGTAGCTG 660
QY 661 GAGAGCTGTGTGCGCGGAGGAGTCAATGCTCAAGTTCCAGCGCGGAGCTTGAACAT 720
DB 661 GAGAGCTGTGTGCGCGGAGGAGTCAATGCTCAAGTTCCAGCGCGGAGCTTGAACAT 720
QY 721 TATCTGCTGAGCGCTGCTGGCAACGCGCGGAGCTTACCGCCCATCCAGCATCTCAAC 780
DB 721 TATCTGCTGAGCGCTGCTGGCAACGCGCGGAGCTTACCGCCCATCCAGCATCTCAAC 780
QY 781 CCCATCAACATGCTGTGTGTCAGGTCGCTTCTTAGAGATCGTACTCGGCGCCCAAG 840
DB 781 CCCATCAACATGCTGTGTGTCAGGTCGCTTCTTAGAGATCGTACTCGGCGCCCAAG 840
QY 841 GTCACCGCAATGCGCGCTGACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTCAAC 900
DB 841 GTCACCGCAATGCGCGCTGACGCTGCGCAACTTCTGTGCTGGCAGAGAAGCTCAAC 900
QY 901 AAAGTGAGTGACAAACACCCCGAGTACTGGGACATGCGATCCCTTCCAGGAGGAG 960
DB 901 AAAGTGAGTGACAAACACCCCGAGTACTGGGACATGCGATCCCTTCCAGGAGGAG 960
QY 961 CTGTGTGAGGACCAACCTGTGACACCTGGCATGGCTGTGATGGGTACCATGTGTAC 1020
DB 961 CTGTGTGAGGACCAACCTGTGACACCTGGCATGGCTGTGATGGGTACCATGTGTAC 1020
QY 1021 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGCTTCCATCAGCTTCCACTGCC 1080
DB 1021 CCCAAGAGAAGCTGCTCTGTCTATTGAGGACGATGGGCTTCCATCAGCTTCCACTGCC 1080
QY 1081 CACGAGCTGGG 1091
DB 1081 CACGAGCTGGG 1091

RESULT 7

AAD35568
ID AAD35568 standard; cDNA; 966 BP.

XX AAD35568;

XX AC

XX 26-JUL-2002 (first entry)

XX Human protease cDNA #1.

XX Human; novel human protein; NHP; protease; biological disorder; obesity;

XX high blood pressure; arthritis; connective tissue disorder; infertility;

XX gene therapy; enzyme; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..966

XX /*tag= a

XX /product= "Human protease #1"

XX WO200226949-A2.

PD	04-APR-2002.	
XX		
PF	27-SEP-2001; 2001WO-US30350.	
XX		
PR	29-SEP-2000; 2000US-236689P.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
XX		
PI	Friddle CJ, Hilbun E;	
XX		
DR	WPI: 2002-372123/40.	
XX		
DR	P-PSDB; AAE22540.	
XX		
PT	Novel nucleic acid encoding a human protease, useful as a hybridization	
PT	probe for screening libraries and assessing gene expression patterns -	
XX		
PS	Disclosure; Page 34; 41pp; English.	
XX		
CC	The present sequence is a cDNA encoding novel human protein (NHP),	
CC	human protease. NHPs share structural similarity with animal proteases	
CC	particularly zinc metalloproteases. Sequences of the invention are	
CC	useful in therapeutic, diagnostic and pharmacogenomic applications.	
CC	NHP polynucleotides are used as hybridisation probes for screening	
CC	libraries and assessing gene expression patterns. They can also be	
CC	used for treating related biological disorders such as obesity, high	
CC	blood pressure, arthritis, connective tissue disorders and infertility.	
CC	They are also used in gene therapy.	
XX		
SQ	Sequence 966 BP; 161 A; 316 C; 305 G; 184 T; 0 Other;	
Query Match 33.6%; Score 959; DB 24; Length 966;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 ATGCTTCTGCTGGGATCTTAACCTGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG	60
DB	1 ATGCTTCTGCTGGGATCTTAACCTGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG	60
QY	61 CCAGAGCGGGAGGTAGTGTCCATCCGACTGACCGCGGACATTAACGGCGCGCTAC	120
DB	61 CCAGAGCGGGAGGTAGTGTCCATCCGACTGACCGCGGACATTAACGGCGCGCTAC	120
QY	121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT	180
DB	121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCAATTTTCAGATCACAGCATTT	180
QY	181 CAGGAGGACTTTTACCTACACTGACCGCGGATCTTCAGTTCTGGCTCCGCTTCTCC	240
DB	181 CAGGAGGACTTTTACCTACACTGACCGCGGATCTTCAGTTCTGGCTCCGCTTCTCC	240
QY	241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTACCGGGGCTTTTCAGACCTGCGACGC	300
DB	241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTACCGGGGCTTTTCAGACCTGCGACGC	300
QY	301 TGCTTCTATTCTGGGAGCTGAACCGCGAGCGCGGACTTCGCTGCTGTGAGCCTGTC	360
DB	301 TGCTTCTATTCTGGGAGCTGAACCGCGAGCGCGGACTTCGCTGCTGTGAGCCTGTC	360
QY	361 GGGGGCTCCCGGAGCTTTGGCTACCGAGCGCGGCTATGTCATTAGCCGCTGCC	420
DB	361 GGGGGCTCCCGGAGCTTTGGCTACCGAGCGCGGCTATGTCATTAGCCGCTGCC	420
QY	421 AATGCTAGCGCGCGGGCGGACGCAACAGCGAGCGCGCACACCTTCTCCAGCGCGG	480
DB	421 AATGCTAGCGCGCGGGCGGACGCAACAGCGAGCGCGCACACCTTCTCCAGCGCGG	480
QY	481 GGTGTTCGGCGGCGCTTCGGAGACCCACCTCTCGTGGGGTGGCTCGGGCTGG	540
DB	481 GGTGTTCGGCGGCGCTTCGGAGACCCACCTCTCGTGGGGTGGCTCGGGCTGG	540
QY	541 AACCCCGCATCTTACCGGCGCTTACAGCCCTTACAGCGCGGAGGGCGGCTTCGGGAG	600
DB	541 AACCCCGCATCTTACCGGCGCTTACAGCCCTTACAGCGCGGAGGGCGGCTTCGGGAG	600

QY	601 AGTCGTAGCCGGCGCAGGCTCTGGCGCGCCAAAGCGTTTCGTCTATCCCGCGTACG	660
DB	601 AGTCGTAGCCGGCGCAGGCTCTGGCGCGCCAAAGCGTTTCGTCTATCCCGCGTACG	660
QY	661 GAGACGCTGGTGTGCGCGGAGTCAATGTCAGTTCCACGGCGGACCTGGACAT	720
DB	661 GAGACGCTGGTGTGCGCGGAGTCAATGTCAGTTCCACGGCGGACCTGGACAT	720
QY	721 TATCTGCTGACGCTGTGGCAACGGCGCGGACTCTTACCGCCATCCAGCATCTCAAC	780
DB	721 TATCTGCTGACGCTGTGGCAACGGCGCGGACTCTTACCGCCATCCAGCATCTCAAC	780
QY	781 CCATCAACATCTGTTGGTCAAGGCTCTCTTTAGAGATCTGACTCCGGGCGGCAAG	840
DB	781 CCATCAACATCTGTTGGTCAAGGCTCTCTTTAGAGATCTGACTCCGGGCGGCAAG	840
QY	841 GTACCCGCAATCGGCGCTGACGCTGCGCAACTTCTGTCCCTGGCAGAGAAGCTGAAC	900
DB	841 GTACCCGCAATCGGCGCTGACGCTGCGCAACTTCTGTCCCTGGCAGAGAAGCTGAAC	900
QY	901 AAAGTGAGTGACAGCAACCCGCGAGTACTGGGACACTGCTCTTCCAGGCGAGGA	959
DB	901 AAAGTGAGTGACAGCAACCCGCGAGTACTGGGACACTGCTCTTCCAGGCGAGGA	959
RESULT 8		
AA95831		
ID	AAA95831 standard; cDNA; 1143 BP.	
XX		
AC	AAA95831;	
XX		
DT	23-FEB-2001 (first entry)	
XX		
DE	Human metalloproteinase ADAMTS-5 cDNA.	
XX		
KW	Human; ADAMTS-5; metalloproteinase; ADAM;	
KW	a disintegrin and metalloproteinase domain; thrombospondin domain;	
KW	vaccine; neutropenic; neuroprotective; antiparkinsonian;	
KW	cerebroprotective; cytosolic; antiarthritic; immunosuppressive;	
KW	Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;	
KW	autoimmune disease; brain tumour; brain injury; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	
FT	CDS	
FT	1..1143	
FT	/tag= a	
FT	/partial	
FT	/product= "ADAMTS-5"	
XX		
PN	WO200053774-A2.	
XX		
PD	14-SEP-2000.	
XX		
PF	08-MAR-2000; 2000WO-US06237.	
XX		
PR	08-MAR-1999; 99US-0264585.	
XX		
PA	(NEUR-) NEUROCRINE BIOSCIENCES INC.	
XX		
PI	Kelner GS, Clark M, Maki RA;	
XX		
DR	WPI: 2000-594326/56.	
DR	P-PSDB; AAB21261.	
XX		
PT	Polynucleotide encoding novel members of a disintegrin,	
PT	metalloproteinase and thrombospondin domain protein family used to	
PT	prevent and treat Alzheimer's disease, cancer and autoimmune diseases -	
XX		
PS	Claim 2; Fig 22; 129pp; English.	
XX		
CC	The present sequence encodes human metalloproteinase ADAMTS-5. The	

tech
No


```

PR 11-NOV-1999; 99JP-0321740.
PR 16-MAY-2000; 2000JP-0144020.
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
DR WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis
XX
PS Example 13; Page 66; 85pp; Japanese.
XX
CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC an adaptor primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 30 BP; 8 A; 6 C; 10 G; 6 T; 0 other;
Query Match 1.1%; Score 30; DB 22; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 TTACTACCTGACGCGGATGCTCAGT 220
DB 30 TTTACCTACCTGACGCGGATGCTCAGT 1
RESULT 11
AAH41030/c
ID AAH41030 standard; DNA; 30 BP.
XX
AC AAH41030;
XX
DT 23-AUG-2001 (first entry)
DE PCR primer specific for metalloprotease DNA isolation.
XX
KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
KW osteopathic; antiarthritic; primer; adaptor; ss.
XX
OS Homo sapiens.
XX
PN WO200134785-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-JP07917.
XX
PR 11-NOV-1999; 99JP-0321740.
PR 16-MAY-2000; 2000JP-0144020.
XX
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
XX
XX WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis
XX
PS Example 13; Page 66; 85pp; Japanese.
XX
CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC an adaptor primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 30 BP; 8 A; 6 C; 10 G; 6 T; 0 other;
Query Match 1.1%; Score 30; DB 22; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 TTACTACCTGACGCGGATGCTCAGT 220
DB 30 TTTACCTACCTGACGCGGATGCTCAGT 1
RESULT 12
AAH41017/c
ID AAH41017 standard; DNA; 41 BP.
XX
AC AAH41017;
XX
DT 23-AUG-2001 (first entry)
DE PCR primer specific for human metalloprotease DNA.
XX
KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
KW osteopathic; antiarthritic; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200134785-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-JP07917.
XX
PR 11-NOV-1999; 99JP-0321740.
PR 16-MAY-2000; 2000JP-0144020.
XX
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
XX
XX WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis
XX
PS Example 4; Page 62; 85pp; Japanese.
XX
CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 41 BP; 5 A; 16 C; 13 G; 7 T; 0 other;
Query Match 1.1%; Score 30; DB 22; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1720 GCCTCCGAAAGAGCTCCGGGAGGAGCAG 1749

```

```

Db 41 GCCTCGGAAGAGCTTCGGGAGGAGCAG 12
|||||
RESULT 13
AAH41018
ID AAH41018 standard; DNA; 27 BP.
XX
AC AAH41018;
XX
DT 23-AUG-2001 (first entry)
XX
DE PCR primer specific for human metalloprotease DNA SEQ ID 9.
XX
KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
KW osteopathic; antiarthritic; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200134785-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-JP07917.
XX
PR 11-NOV-1999; 99JP-0321740.
XX
PR 16-MAY-2000; 2000JP-0144020.
XX
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
XX
DR WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis -
XX
PS Example 5; Page 63; 85pp; Japanese.
XX
CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 27 BP; 5 A; 6 C; 11 G; 5 T; 0 other;
Query Match 0.9%; Score 27; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1534 AAGCACAGGTGGATGTTCTCTGGGCC 1560
|||||
Db 1 AAGCACAGGTGGATGTTCTCTGGGCC 27

RESULT 14
AAH41019/c
ID AAH41019 standard; DNA; 37 BP.
XX
AC AAH41019;
XX
DT 23-AUG-2001 (first entry)
XX
DE PCR primer specific for human metalloprotease DNA SEQ ID 10.
XX
KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
KW osteopathic; antiarthritic; PCR primer; ss.

```

```

XX Homo sapiens.
OS WO200134785-A1.
PN 17-MAY-2001.
PD 10-NOV-2000; 2000WO-JP07917.
PF 11-NOV-1999; 99JP-0321740.
PR 16-MAY-2000; 2000JP-0144020.
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;
XX
DR WPI; 2001-343602/36.
XX
PT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis -
XX
PS Example 5; Page 63; 85pp; Japanese.
XX
CC This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 37 BP; 7 A; 14 C; 14 G; 2 T; 0 other;
Query Match 0.9%; Score 27; DB 22; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2824 CTGGACTTCTCGTCTCGAGCCGTGC 2850
|||||
Db 37 CTGGACTTCTCGTCTCGAGCCGTGC 11

RESULT 15
AAH41024/c
ID AAH41024 standard; DNA; 38 BP.
XX
AC AAH41024;
XX
DT 23-AUG-2001 (first entry)
XX
DE PCR primer specific for metalloprotease related DNA.
XX
KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;
KW osteopathic; antiarthritic; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200134785-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-JP07917.
XX
PR 11-NOV-1999; 99JP-0321740.
XX
PR 16-MAY-2000; 2000JP-0144020.
XX
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (KAZU-) KAZUSA DNA RES INST.
XX
PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

```

```
XX WI; 2001-343602/36.
XX DR
XX PT
XX TT Metalloprotease with aggrecanase activity for treating joint diseases
PT especially osteoarthritis -
XX
XX Example 9; Page 65; 85pp; Japanese.
PS
XX This invention relates to a metalloprotease with aggrecanase activity.
CC The invention includes protein and DNA sequences of the metalloprotease,
CC vectors containing the DNA, host cells transformed by the vectors, and
CC antibodies directed against the metalloprotease. The antibodies, protein
CC and DNA sequences can be used in the treatment and prevention of joint
CC diseases, particularly osteoarthritis. The treatment may result in
CC osteopathic and antiarthritic activity. The present sequence represents
CC a PCR primer used in the isolation and characterisation of the
CC metalloprotease gene of the invention.
XX
SQ Sequence 38 BP; 6 A; 13 C; 8 G; 11 T; 0 other;

Query Watch          0.98; Score 27; DB 22; Length 38;
Best Local Similarity 100.0%; Pred.No. 0.043;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2035 AATAAGAGCTGCAAGAAGGTGACTGGA 2061
      |||||||
DB 38 AATAAGAGCTGCAAGAAGGTGACTGGA 12

Search completed: May 16, 2003, 01:01:50
Job time : 587 secs
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:21:31 ; Search time 29 Seconds
(without alignments)
963.854 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLLGILTLAFAGRTAGGSE.....DQCNLHRKPOHDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2486.5	48.2	967	4	US-09-130-491-2
2	2115	41.0	905	4	US-09-369-364A-9
3	1915	37.1	837	4	US-09-122-126B-2
4	1900	36.8	608	4	US-09-130-491-13
5	1854	35.9	930	4	US-09-369-364A-2
6	1848.5	35.8	1882	4	US-09-369-364A-13
7	1847.5	35.8	930	4	US-09-122-126B-15
8	1798.5	34.8	874	4	US-09-369-364A-15
9	1607	31.1	551	4	US-09-130-491-16
10	1273	24.7	481	4	US-09-130-491-8
11	1264	24.5	518	4	US-09-369-364A-22
12	1248.5	24.2	1224	4	US-09-930-872-4
13	1218	23.6	1211	4	US-09-491-522-5
14	1205.5	23.4	997	4	US-09-369-364A-7
15	1152	22.3	1081	4	US-09-369-364A-17
16	1144	22.2	1205	4	US-09-491-522-11
17	1035	20.1	859	4	US-09-369-364A-5
18	662.5	12.8	245	4	US-09-369-364A-11
19	577.5	11.2	566	4	US-09-491-522-7
20	550.5	10.7	525	4	US-09-369-364A-2h
21	441	8.5	491	4	US-09-930-872-2
22	380.5	7.4	450	4	US-09-369-364A-19
23	349	6.8	812	4	US-09-632-098-4
24	339	6.6	802	4	US-09-632-098-2
25	306	5.9	464	4	US-09-411-329C-14
26	305.5	5.9	616	4	US-09-608-790-1
27	304	5.9	462	4	US-09-411-329C-3

28	304	5.9	462	4	US-09-411-329C-17	Sequence 17, Appl
29	301	5.8	529	2	US-08-836-442-3	Sequence 3, Appl
30	288	5.6	621	4	US-09-026-001A-6	Sequence 6, Appl
31	285	5.5	855	4	US-09-813-819-2	Sequence 2, Appl
32	285	5.5	855	4	US-09-920-048-2	Sequence 2, Appl
33	281	5.4	621	4	US-09-026-001A-18	Sequence 18, Appl
34	279	5.4	769	1	US-08-243-542-4	Sequence 4, Appl
35	279	5.4	769	1	US-08-477-407-4	Sequence 4, Appl
36	279	5.4	769	1	US-08-484-355-4	Sequence 4, Appl
37	277.5	5.4	814	4	US-09-813-819-4	Sequence 4, Appl
38	277.5	5.4	814	4	US-09-920-048-4	Sequence 4, Appl
39	276	5.3	613	4	US-09-026-001A-10	Sequence 10, Appl
40	274.5	5.3	751	2	US-08-836-443-3	Sequence 3, Appl
41	263	5.1	592	4	US-09-026-001A-14	Sequence 14, Appl
42	260	5.0	670	1	US-08-243-542-3	Sequence 3, Appl
43	260	5.0	670	1	US-08-477-407-3	Sequence 3, Appl
44	260	5.0	670	1	US-08-484-355-3	Sequence 3, Appl
45	248.5	4.8	1170	1	US-08-313-288B-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-130-491-2
; Sequence 2, Application US/09130491
; Patent No. 6418974

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

; FILE REFERENCE: 09404/041001

; CURRENT APPLICATION NUMBER: US/09/130,491

; EARLIER FILING DATE: 1998-08-07

; EARLIER FILING DATE: 1997-09-05

; EARLIER APPLICATION NUMBER: US 60/058,108

; EARLIER FILING DATE: 1997-08-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 967

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-130-491-2

Query Match 48.2%; Score 2486.5; DB 4; Length 967;
Best Local Similarity 49.0%; Pred. No. 1.5e-198;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

QY	1	MLLLGILTLAFAGRTAGGSEPREVVVPIRLDPDINGRRYYWRGPEDSGDGLIFQITAF	60
DB	36	LLLLAALLAVSALGRPSEDEELVVP-ELE-----RAP---GHGTTTLRLHAF	81
QY	61	QEDFYLHLTPDAQFLAPAFSTEHLG-----VPLQGITGSSDLRCFCYSGDVNAEPDSF	114
DB	82	DQQLDLRLPDSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA	136
QY	115	AAYSLCGGLRGARYGAEVVISPLPNAS---APAAQRNSQGA-----HLLQ---RGVPGG	165
DB	137	AALSICGVRGAFYLLGEAYFIQPLPAASERLATAPEKPPAPLQFHLRLKRNQDVG	196
QY	166	PSG-----DPTSRC-----GVASG--WNPAIRALDPYKPRRAGFESKRRRS	207
DB	197	TCGVVDEPRPTKATEDEDETEDETEDETEDETEDETEDETEDETEDETEDETEDETE	247
QY	208	GRAKREVSIPRYVETLVVADESMVKFHGADLEHYLLTLTALTAARLYRHPISILNPINIVV	267
DB	248	IRKKEVSSHRYVETMLVADQSMARFHSGLKHYLLTLFSAARLYKHPISIRNSVSLVV	307
QY	268	KVLLLRDRSGPKVTGNAALTALNFCAWOKKLNKVDKHPYWDRTAILFTRODLGATTC	327
DB	308	KILVHDEQKGPVTSNAALTALNFCAWOKKLNKVDKHPYWDRTAILFTRODLGATTC	367

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 837

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-122-126B-2

Query Match

Best Local Similarity 37.1%; Score 1915; DB 4; Length 837;

Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

```

1 MLLGLTILAFAGRTAGGSEPREVVVPIRLDPDINGRYYWRGPEDSGQSLFIQTAF 60
37 LLLLLLALPSARLASPLPREEIIVPEKLGSLV-----PGSGAPARLLCKLQAF 88
61 QEDFYLHLPDQAFLAPAFSTHLPVQLQGLGGSDLLRRCFYSGDVNAEPLDSFAVSLC 120
89 GETLLELEQDSGVQEGVTUQLQAPB-LUGGAEF--GYLTGTINLLEHESVASLHMD 145
121 GG-LRGAFYGAEYVISPLNAPAAQNSQGAHLQRRGVPGGPSGLPTSRCCVASG 179
146 GGALLGVLYRGAELHLQPLEGGTPNSA--GGPGAHLRRK----SPASQGPNCNV--- 196
180 WNPAILRALDPYKPRAGGEGSRRRSGRAKRFVSIPIRYVETLVVADESVMKPHGADLE 239
197 -----KAPLGSPSPRR--RAKRFASLSRFVETLVVADDDKMAAHPGAGLK 239
240 HYLLTLLATAARLYRHPSTILNPINWVVKVLLLRDRDGGPKVTGNAALTLENECAWQKL 299
240 RYLLVTMAAAKAFKHPSTRNPVSLVTVLILSGEGEPQVGPSSAQTIRSFCAWQRL 299
300 NKVSKHPHYDWTALIFTRQDLGGATCTDTLGMADVTGMDPKRSCSVIADGLPSAFTT 359
300 NTPEDSDPHEDTALIFTRQDLGGVSTCTDTLGMADVTGMDPARSCAIVEDDGLQSAFTA 359
360 AHELGHVFNMPHDNVKCEEEVFKL--RANHMSPTLIQIDRANPWSACSAAIITDFDLSG 418
360 AHELGHVFNMLHDNKKPCISLNGPLUSTSRHVNAPVMAHVDPPEWPSCSAFIFDLDNG 419
419 HGCDLLDQSPKPISLPEDLGASTLSQOCELAFAFGVSGKPCPYM--QYCTKLWCTGKAKGQ 477
420 YGHCLLDKPEAPLHPVTFPGKYDADRCQQLITFGPDSRHCPQLPPCAALWCSGHLNGH 479
478 MYCQTRHPHWADGTCGEGEKLKLGACVVERHNLNKHVRD--GSAKWDPYGPSCRTCGG 535
480 AMCQTKHSPWADGTCGPAQACMGRCCLHMDQLQDFNIPQAGGWPMPAGDCSRTCGG 539
536 VOLARRQCTNPTPANGKYCYGVRKYRSCNLEPCPSSASGKSPREOCEAFNGYNHSTN 595
540 VQFSSRCDTRPVNPGKYCEGRTRFRSCNTEDCP--TGSALTREBOCA--YINHTD 595
596 RLTL---AVAWVPKYSGVSPRDKCKLICRANGTGYFYVYAPKVVVDGTLCPDSTSVQVQ 652
596 LFKSPGPMWVPRYTVGAPDQCKLTQARALGYVYVLEPRVVDGTPCSPDSSVCVQ 655
653 KCIKAGCDNGLSKRDKCVCVGGDNKSKKVTGLTKPMHGFNVVAIPAGASSIDIR 712
656 KCIHAGCDRIIGSKKKDKCMKCGDGGCSKQSGSKGSKFRKFRYGNVVTIPAGATHILVR 715
713 ORGYKGLIGDNNYLALNKSQKYLINGHFFVYSAVERDLVVKGSF--LRYSGTGTAVESLQA 771
716 QOQNGP--HRSIYLALKLIPDGSYALNGEYTLMPSTDVLPAGVSLRYSATASETLSG 773
772 SRPILPULTEVLVSGMKTPPRVRYSYFLPKREPKDSSHPKD 814
774 HGPLAQPLTLQVLVAGNPQDTRLRYSEFFVPRPTSTPRPTPD 816
```

RESULT 4

US-09-130-491-13

; Sequence 13, Application US/09130491

; Patent No. 6416974

; GENERAL INFORMATION:

```

; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130.491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-130-491-13
```

Query Match 36.8%; Score 1900; DB 4; Length 608;

Best Local Similarity 49.3%; Pred. No. 6e-150;

Matches 339; Conservative 114; Mismatches 146; Indels 88; Caps 8;

```

269 VLLLRDRSGPKVTGNAAALTLRNFCAWKKLNKVDKHPHYDWTALIFTRQDLGGATCTD 328
4 ILVHDEQKGPVETSNAALTLENECAWQKQHNPPSDRAEHYDATALIFTRQDLGGATCTD 63
329 TLGMADVTGMDPKRSCSVIETDDGLPSAFTTAHELGHVFNMPHDNVKCEEEVFKLRANH 388
64 TLGMADVTGMDPKRSCSVIETDDGLQAAFTTAHELGHVFNMPHDNADKACASLNGVNOJSH 123
389 MMSPTLLQIDRANPWSACSAAIITDFDLSGHCGLDLPSPKISLPEDLPGASTYLSQ 448
124 MMASMLNLDHSPWSPSCSAIYMTISFLDNHGGECMLDKPQNPQLPDGLPGTSYDANRQC 183
449 ELAFVGSKPCP--YMYCTKLWCTGKAGQMVQCTRHFPWADGTCGEGKLCIKLGACVER 507
184 QFTFGEDSKHCPDAASTCTLWCTGSGVLVQCTHFPWADGTCGEGKWCINGKCVNK 243
508 HNLNKH--RVDGSAKWDPYGPSCRTCGGVQLARQCTNPTPANGKYCEGVRYKYRS 564
244 TD-RKHFDTPFHGSGWGMWGPWDCSRTCGGVQYTMRECDNPVPKNGGKYCEGKRVYRS 302
565 CNLEPCPSSASGSKFREEOCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANG 624
303 CNLEDCPDN--NGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAG 361
625 TGYFYVYAPKVVVDGTLCPDSTSVQVQKCIKAGCDNGLSKRDKCVCVGGDNKSCXK 684
362 IGYFFVLPKVVVDGTLCPDSTSVQVQKCIKAGCDRIIDSKKKDKCGVGGNGSTCKX 421
685 VTGLTKPMHGYVNVVAIPAGASSIDIRORYKGLIGDNNYLALNKSQKYLINGHFFVYS 744
422 ISSVTSKAPGYHDIITIPIGATNIEVKORNGSRNNGSFLAKAADGTIYLNGDYILS 481
745 AVERDLVWKSLLRYSGTGTAVESLQAASRPILPTEVLVSGMKTPPRVRYSYFLPKPEP 804
482 TLEQDITWYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGNALRPKIKTYFYVKKX- 540
805 REDKSHHPKDPGCPVSLHNSVLSLQVQEPDPRPARVAGSWGSPCSASCGSLQKFAV 864
541 ---KES-----FNAITFS----- 551
865 DCRGSAGQRTVPACDAHHRPVETQACGEPCTWELSAWSPCSKSGCGFORSLKCVGJG 924
552 -----AWIEEGECSKTCGKYKRSKLSCLSD 580
925 GRLLARDQCNLHRKPOE--LDFCVLRPC 950
581 GGVLSHESCDPLKKPKHFDICTMAEC 607
```

RESULT 5

US-09-369-364A-2

; Sequence 2, Application US/09369364A

; Patent No. 6391610

; GENERAL INFORMATION:

; APPLICANT: Apte, Suneel

; APPLICANT: Hurskainen, Tiina L.

; APPLICANT: Hirohata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

; FILE REFERENCE: 26473/4007/10-30-00

; CURRENT APPLICATION NUMBER: US/09/369,364A

; CURRENT FILING DATE: 1999-08-06

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 930

; TYPE: PRT

; ORGANISM: mus musculus ADAMTS-5

; US-09-369-364A-2

Query Match 35.9%; Score 1854; DB 4; Length 930;

Best Local Similarity 43.1%; Pred. No. 7.6e-146;

Matches 369; Conservative 116; Mismatches 265; Indels 106; Gaps 13;

Qy 90 GLTGSSDLRRCFYSGVDNAPDSFAVSLCGGLRGAFGRGAEYVLSLPLNASAPAAQR 149

Db 121 GLSASSGRHGCFYRGTVGDSRSLAVFLDGLGDFFAVKHARYTLKPLRGSWAEYR 180

Qy 150 -----NSOGAHLLORRG-----VPGSGDPTSRGVSAGWNPAILRALDPY 191

Db 181 IYDGSRLHYVNRGFSFALPPRASCETPASPG----- 217

Qy 192 KPRRAGEFSRRRSRG-----RAKRFVSIPIRYVETILVAD 227

Db 218 -PQESPVSVHSRRRSALAPQLLDHSAFSPGNAGPQTWNRRRRSISRARQVELLIVAD 276

Qy 228 ESMVKFHGADLEHLLVILLATARLYRHPSTILNPIVIVVYKVLRLDRDSGPKVTGNAAL 287

Db 277 SSMARMYGRGLQHLVLLMASTANLRSHASTENHRLAVVYKVVLLDKDTSLEVSNAAT 336

Qy 288 TLRNFCWOKLNKVSQKHPYDWTALFTRODLGATTCDTLGMADVGTMCDPKRSCSV 347

Db 337 TLNFKCKWOHNLGDDHEHYDAALFTREDLCGHSHCDTLGMADVGITCSPERSCAV 396

Qy 348 IEDDGLPSAQTAAHELGHVFNHNDNVKVEEVEGKLRANHMSPTLIQIDRANPWSACS 407

Db 397 IEDDGLHAAFTVAHEIGLLGLSHDDSKFCBENFGTTEDKRLMSSILTSIDASKPWSKCT 456

Qy 408 AALITDFLDGSHGCDLLDQSPKISLPLDGPASYTLISQCELAFAFGVSKICPYMQVCTK 467

Db 457 SATITEFLDDGHGNCLLDLPKQILGPEELPGQYDQICNLTFGPEYVVCFGMDVCAR 516

Qy 468 LWCTGKARKQMVCTRHFPWADGTSCGEGKLCGLKACVER-----HNLNKHVRVDGSWAKW 522

Db 517 LWCAVVRGQMVCTLTKLPVAVGTCGKGRVCLQCKVDKTKKYYSTSH---GNMGSW 573

Qy 523 DYPGCSFTCGGQVLARRQCTNTPANGKYCEGVKVRKTRSCNLEPCPSASCKSREE 582

Db 574 GPWGQSRCSGGVQGFAYRHCNAPRNSRGYCTGKRAIVRSCSVTPCP--PNGKSPRHE 631

Qy 583 QCEAFGNHSTNRLTLAVAMPYKSGVSPRDKCKLICRANGTCGYFVLAPKVVVDGTCLS 642

Db 632 QCEAKNGVQSDAKGVKTFVWPKYAGVLPADVCKLICRAGTKGYVVPKVTIDGTGECR 691

Qy 643 PDSTSVCCVQGCICIKAGCDNGLSKRRFDKCGVCGGDNKSKKVTGLFTKPMHGYNFVAI 702

Db 692 PYSNVCVIRGRCVTRGCDGIIGLSQYDKCGVCGGDNSSCTKIIGTFNKKSQKGYTDVRI 751

Qy 703 PAGASSIDIRQYKGLIGDDNYLALKNSQCKYLLNGHFVVSAYERDLVVKGLLRYSGT 762

Db 752 PEGATHIKVROFKADQDTRFPAYLALKKKTGTEYLINGKYMISTSETIIDINGTWNVNSGW 811

Qy 763 GRAVESL-----QASRPITLPTVEVLSVGRKMTPPRVRYSFYLPKPRELKS---SHPKD 814

Db 812 SHRDQFLHGMYSATKEI---LIQVILATDPTKALGVRYSFVFPKKTQKVNSVISHGSN 868

Qy 815 PRGPSVLHNSVLSLSNOVEQPDPRPPARWAGSWGPGSCSGGLQKRAVDCRGSAGQRT 874

Db 869 KVGPP---HSTQL-----QWVTGPMWLAACSRCTDCTGWHTRTVQCC-DGNRKL 909

Qy 875 VPACDAAHRPVETQAC 890

Db 910 AKGCLLSQRPSAFKQC 925

RESULT 6

US-09-369-364A-13

; Sequence 13, Application US/09369364A

; Patent No. 6391610

; GENERAL INFORMATION:

; APPLICANT: Apte, Suneel

; APPLICANT: Hurskainen, Tiina L.

; APPLICANT: Hirohata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

; FILE REFERENCE: 26473/4007/10-30-00

; CURRENT APPLICATION NUMBER: US/09/369,364A

; CURRENT FILING DATE: 1999-08-06

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 1882

; TYPE: PRT

; ORGANISM: Homo sapiens ADAMTS-9

; FEATURE:

; NAME/KEY: MOD_RES

; LOCATION: (468)

; OTHER INFORMATION: Xaa = C

; NAME/KEY: MOD_RES

; LOCATION: (521)

; OTHER INFORMATION: Xaa = Y

US-09-369-364A-13

Query Match

Best Local Similarity 35.8%; Score 1848.5; DB 4; Length 1882;

Matches 396; Conservative 144; Mismatches 340; Indels 121; Gaps 23;

Qy 22 EREVVVPIRLDP-----DINGRRYYW-----RGPEDSDGQGLIFQITAFQ 61

Db 49 EYEIVSPIKVNALGPEFFPTNVHFKTRRSINSATDPWAFASSSSSTSSQAHYRLSAFG 108

Qy 62 EDFYLLHTLTPDAQFLAPAFSTEHLGVP-----LQGLGGSSDLRRCFYSGVDNAPDSFRAV 117

Db 109 QQFLFNLTNAGAFIAPLFTVTLTGTPGVNQTIFYSEEAELKHCFYK----- 155

Qy 118 SLCCGLRGAFGRGAEYVLSPLPNASAPAAQNSOGAHLLORRGVPGSGDPTSRGVA 177

Db 156 RLCQVQLRAHG-----RHQPLLREHNRHSKDKKTRAKWGERINLADGV 202

Qy 178 SGWNPAL-LRALDPYKPRRAGEFSRRRSRGRKREVSIPRYVETLVVADSWMYKFHGA 236

Db 203 AALNSGLATEAASAYGNKTDNTRKTRHRT---KRELSYPRFEVLVVDNRNMYSYHGE 259

Qy 237 DUEHYLLTLTAARLYRHPSTILNPIVIVVYKVLRLDRDSGPKVTGNAALTLRNFCAWQ 296

Db 260 NLQHYILTLMSIVASIKDPSIGNLINIVNLVHNEQDGPDSISFNAAQTTLKNFCOMQ 319

Qy 297 KKLNVKSDKHPEVDWTAILFTRODLGCA--TTCDTLGMADVGTWCDPKRSCSVIEDDGLPS 355

Db 320 HSNPSGGIIH---DPAVLLTRQDICRAIDKCDTGLGAEGLTICDIPYRSCSISEDSGLST 375

Qy 356 AFTTAHELGHVFNHNDNVKVEEVEGKLRANHMSPTLIQIDRANPWSACSAAITDPL 415

Db 376 AFTIAHELGHVFNHNDNVKVEEVEGKLRANHMSPTLIQIDRANPWSACSAAITDPL 434

Qy 416 DSHGBCDCLLDOP--SKFISLIPEDLPGASYTLISQCELAFAFGVSKICPYMQVCTKLCW--T 471

Db 435 DTGYGECLLNEPESRYPILPVQLPGLILYNNKQXELIFGFSOVCPYMMQCRRLNCNNVN 494


```
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-8

Query Match
Best Local Similarity 24.7%; Score 1273; DB 4; Length 481;
Matches 236; Conservative 87; Mismatches 141; Indels 38; Gaps 12;

QY 402 PWSACAAIITDLSGSHGDCLLDQSPKIPSLPDLFCASYTLSSQOCELAFYGVSKPCP- 460
Db 7 PWSPCSAVYLTLLDDGHGDCLLD-----DGHSTLYELDQCKQIFGPDFRHCNP 56

QY 461 --YMOYCYKLMCTGKRAKQWVCOTRH--FPWADGTCGEGKLCIKGACVFNHNLN--KHR 514
Db 57 TSVEDICVOLWCRHRSDSEPICTHKNASLLWADGTPCGHLCDLGDGSCVIREVENPKAV 116

QY 515 VDSWAKWDPYGPCSRTCGGGVQLARRCTNPTPANGKGYCEGVVRYRSCNLEPCPSSA 574
Db 117 VDGWGPWGPWQCSCRTCGGGIQFSNRECDNPAPQNGRFLGERVKYQCKTECP--P 174

QY 575 SKSSTREDOCEAFNGYNHSTNRLTLAVAWVPKYSGVSPDRCKKLCIRANGTGYFVVLAPK 634
Db 175 NGKSFREDOCEKYNAYNH--TLDGDNFLQWVPKYSGVSPDRCKLFCRAR--KSEFKVFETK 233

QY 635 VVDGTLCSPDSTSVCOVKCKTAKAGDGNLGSKKRFDKCGVCGDNKCKKAVTGLTKPMH 694
Db 234 VIDGTLGPDTLAICVRGQCVKAGCDHVNSPKKLDKCGVCGGKTACRKGVSFTFSY 293

QY 695 GYNFVVAIPAGASSIDIRQYKGLIGDDNYLALKNQSGKYLLNGHFWGSAVERDLVVKG 754
Db 294 GYNDIVTIPAGATNIDVQRSPGQVNDGYSYALKTANGQVLLNGNLAIENAEQDILMKG 353

QY 755 SLLRYSGTGTAVESLQASRPILPTVEVLSV--GKMTPPRVRYSFYLPKPEPKEDKSSHPK 813
Db 354 TILKYSGSWATLERLQSQALPELITVOLLIVSGEVFPKVKYFFVFNPE--TDFNVQSS 411

QY 814 DPRGFSVLHNSVLSNSQVQPDPRPARWAGSWGPCSCSGSLQKRAVDCRGSGAGOR 873
Db 412 KERAST--NIITSL-----PYAEWVLGDWSECPSTCGGGWORRIVECRDPDSQA 458

QY 874 TVPACDAHRPVETOACG-EPC 894
Db 459 S-DTCDALKPEDARPCGSQPC 479

RESULT 11
US-09-369-364A-22
; Sequence 22, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 22
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-5
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (99)
; OTHER INFORMATION: Xaa = M
US-09-369-364A-22

Query Match
Best Local Similarity 24.5%; Score 1264; DB 4; Length 518;
Matches 244; Conservative 73; Mismatches 185; Indels 36; Gaps 10;

; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-8

Query Match
Best Local Similarity 24.7%; Score 1273; DB 4; Length 481;
Matches 236; Conservative 87; Mismatches 141; Indels 38; Gaps 12;

QY 402 PWSACAAIITDLSGSHGDCLLDQSPKIPSLPDLFCASYTLSSQOCELAFYGVSKPCP- 460
Db 7 PWSPCSAVYLTLLDDGHGDCLLD-----DGHSTLYELDQCKQIFGPDFRHCNP 56

QY 461 --YMOYCYKLMCTGKRAKQWVCOTRH--FPWADGTCGEGKLCIKGACVFNHNLN--KHR 514
Db 57 TSVEDICVOLWCRHRSDSEPICTHKNASLLWADGTPCGHLCDLGDGSCVIREVENPKAV 116

QY 515 VDSWAKWDPYGPCSRTCGGGVQLARRCTNPTPANGKGYCEGVVRYRSCNLEPCPSSA 574
Db 117 VDGWGPWGPWQCSCRTCGGGIQFSNRECDNPAPQNGRFLGERVKYQCKTECP--P 174

QY 575 SKSSTREDOCEAFNGYNHSTNRLTLAVAWVPKYSGVSPDRCKKLCIRANGTGYFVVLAPK 634
Db 175 NGKSFREDOCEKYNAYNH--TLDGDNFLQWVPKYSGVSPDRCKLFCRAR--KSEFKVFETK 233

QY 635 VVDGTLCSPDSTSVCOVKCKTAKAGDGNLGSKKRFDKCGVCGDNKCKKAVTGLTKPMH 694
Db 234 VIDGTLGPDTLAICVRGQCVKAGCDHVNSPKKLDKCGVCGGKTACRKGVSFTFSY 293

QY 695 GYNFVVAIPAGASSIDIRQYKGLIGDDNYLALKNQSGKYLLNGHFWGSAVERDLVVKG 754
Db 294 GYNDIVTIPAGATNIDVQRSPGQVNDGYSYALKTANGQVLLNGNLAIENAEQDILMKG 353

QY 755 SLLRYSGTGTAVESLQASRPILPTVEVLSV--GKMTPPRVRYSFYLPKPEPKEDKSSHPK 813
Db 354 TILKYSGSWATLERLQSQALPELITVOLLIVSGEVFPKVKYFFVFNPE--TDFNVQSS 411

QY 814 DPRGFSVLHNSVLSNSQVQPDPRPARWAGSWGPCSCSGSLQKRAVDCRGSGAGOR 873
Db 412 KERAST--NIITSL-----PYAEWVLGDWSECPSTCGGGWORRIVECRDPDSQA 458

QY 874 TVPACDAHRPVETOACG-EPC 894
Db 459 S-DTCDALKPEDARPCGSQPC 479

RESULT 12
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match
Best Local Similarity 32.9%; Score 1248.5; DB 4; Length 1224;
Matches 332; Conservative 126; Mismatches 386; Indels 165; Gaps 38;

QY 51 OGLIFQITAFQEDFYHLHTLTPDAQFLAPAFSTEHGLGVPLQGLTGGSS-----DLRRCFY 103
Db 95 ESLHLRLKGRHDFHDLRTSSSLVAPGFIVQTL-----GKTGKSVQTLPPDEF--CFY 147

QY 104 SGDVAEPDPSFAANSLCGLRGAFYRGAEVYISPLPNASAPAAQNSQG----AHLQRR 160
Db 148 QGSLSHRNSVALSTCQGLSMIRTEADYFLRPLPSHLKGLRAQAGSSPSHVLYKR 207

QY 161 G-----VPGGSGDPTSR-----CGVASGNWPA-----ILRA 187
Db 208 STEPIAPGASEVLVTSRTELAAHQPLHSSDLRLGLPQKQHFGRKKYMPQPPKEDFIL 267
```

```
QY 188 LDPYPRRAGFESRRSRRAKRFVSPRIVETLVVADESVMKFGH-ALEHLYLLTLL 246
Db 268 PDEYK---SCLHRKSLLRSHNEEL-----NVEVLVVVDKMMQNHGHEINITTYVIL 319
QY 247 ATAARLYRHPSTILNINIVVVLRLDRDGPVKYTGNAALTRNFCAWOKKL-NKVSDBK 305
Db 320 NMSVALFKDGTGGNINIAIVGLIIEDEQPLVISHADHTLSFCQWLSLMCKDKDTR 379
QY 306 HPEYWDTAILFTRODLG--ATTCDTLGMADVGTWCDPKRSVIEDDGLPSAFTTAHEL 363
Db 380 H---DHAILLTGIDCSWKNPCDGLFAPISGMCKSYRSCCTINEDTGLGLAFTIAHES 435
QY 364 GHVFNPHDNV-KVCEEVFGKLRANHMSPTLIOIDRANPWSACSAIIIDFLDSGHGDC 422
Db 436 GHNFMIHDGEGMCKKSG-----NIMSPTLAGNNGVFSWSPCSRQYLRHFLSTAQAIC 490
QY 423 LLDPSKPI---SLPDLPGASYTILSQOCELAFGVGSKPCPY---MOYCIKLMC--TKA 474
Db 491 LADOP-KPVKEYKYPEKPLGELYDANTOCKWQFGEKAKLCMLDKFKIDCKALWCHRIGRK 549
QY 475 KGMVQOTRHPWADTSGEGKCLKGACVERHNLNKHRVDGSHAKWDFYQPCSRTCGG 534
Db 550 -----CETKFMPEAAETICGDMWCGGQCVKYGDEGPKPTHGHSWDSWSNSPCSRTCGG 604
QY 535 GVOLARROCTNPTTPANGKYCEGVRYKYSRNCNLEPCSSASGSKSFREQ-----EAFNG 589
Db 605 GVSHRSLCTNPKPSIGGFCGEGSTFKLNSORCPDSD--VDFRAACAEHNSRRFRG 662
QY 590 YNHSTNRLTLAVAWPKYSGVSPDRCKLICRANGTGYFYLAPKVVDGILCSPDSTSV 649
Db 663 RHYK-----WKP-YTQVEDODLCKLYCIAEGDFEFSLSNKVYKDGTPGSEDSRNV 712
QY 650 VQKCKIAGCDNGLSKKRFDCGCGDNKSKKVTGLTKPMH--GYAFVVAIPAGAS 707
Db 713 IDGICERVGCDNLGSDAEDVCGVCGNNSACTIIRGLYTHHHHTNQYHHWYTPSGAR 772
QY 708 SIDRORGYKGLIGDNNYLALKNSQKYLINGHFVVSAYERDLVVKGSLRYSGTGATVE 767
Db 773 SIRIYENV-----STSYISVRNALRRYLLNGHWTVDMPCR-YKFSGTTEDYRRSYNEPE 826
QY 768 SLQASRILEPLAVEVLSVKQWTPPRVYSFYLPKPREDKSHKPDREPS-----819
Db 827 NLIATGPTNETLIVELLFOGR--NPGWAYEYSMPR-----LGTEKQPPACPSYTWAIYRS 879
QY 820 -----VLHNSVLSLSNOVE---QPDPRP-----DARNVAGSWGPC 851
Db 880 ECSVSCGGQMTVREGCYRDLKTFQVNMFCNPKTRPVTVGLVPCVKVACPESNVGNWSAC 939
QY 852 SASCGSLQKRAVDC--RGSAGORTVPA--CDAAHPRVETOAC-GBPC-FTWELSAWSPC 905
Db 940 SRTCGGAQSRPVQCTRRVHYDSEVPASLCPQP-APSSRQACNSQSCPEANSAGPWAEC 998
QY 906 SKSCRGFORSLKCVCHG-----GRLLARQCNLHRKPOELDFCVLRPC 950
Db 999 SHTCGKWRKRAVACKSTNPASRAQLLPDAVCTSEPKPRMHEACLLQRC 1047

RESULT 13
US-09-491-522-5
; Sequence 5, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Lapiere, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
```


US-09-369-364A-17

Search completed: May 9, 2003, 15:24:43
Job time : 36 secs

Query Match 22.3%; Score 1152; DB 4; Length 1081;
Best Local Similarity 31.5%; Pred. No. 3.1e-87;
Matches 318; Conservative 120; Mismatches 400; Indels 172; Gaps 35;

Qy	64	FYLHLPDAQFLAPAFSTHGLGVPLQGLTGGSDLRRCFYSGDVNAE-PTGFAAVSLCGG	122
Db	65	FLNLTRSRLLAGRVSVYEW--TREGLAQWRAARPHCLYAGHLOGQAS\$SHVAISTCGG	122
Qy	123	LRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGGSDPTSRGCV--ASGW	180
Db	123	LHGLIIVADEEYLIPLHGGPKGSRSPESGPHCVVKRSLRHPHLD--IACGVREKFPW	180
Qy	181	--NPAILRALDPYKPRAGFGESRRRRSRRAKRFVSPRYVETLVVADE\$MVKFEG-AD	237
Db	181	KGRPWWLRLTKP--PPARPLGNETERGQPG-LKRSVSRERYVETMDVADKMMVAYHGRD	237
Qy	238	LEHYLLTLATAARLYRHP\$ILNPIVIVVKKVLLLRDRSDGPKVTGNAALILRNECAMOK	297
Db	238	VEQYVLAIMNIVAKLFQDSSLSG\$TNILVTRLILTEDQPTLEITHHAGKSLDSFCKWCK	297
Qy	298	KL-----NKVSDKHPEWDTAILTTRDLC--GATTCTGLMADVGT\$CDPKRSCSVI	348
Db	298	SIVNHSGHGNAIPENGVAHNDTAVLITRYDICIYKKNKPGCTILGLAR-WAECV\$AREAAAS	356
Qy	349	EDGLPSAFTTAHELGHVFNPHDNYKVCEEVFG-----KLRAHHM\$MPTLIQIDRANP--	402
Db	357	MRTLAATSVHHCHEIGTFGMNHDGVGNSGGARGQDPAKLMAHITMKT-----NP\$V	409
Qy	403	WSACSAAIITFDLSDSHGDCLLDOP\$K-PISLPEDLP\$GASYTL\$OQCELAF\$GVG\$KPCPY	461
Db	410	WSSCNRDYIT\$FLD\$GLGLCLNNRPPQD\$FVYPTVAPGQAYDADEQCR\$H\$GVK\$RQCKY	469
Qy	462	MOYCTKLWCTGKAKQWCVQTRH\$P\$WADGT\$CGEGL-----CLKGACV\$ERHNLNKH\$VDG	517
Db	470	GEVSELWCL\$K\$NR--CITNSIPAAEGTLQO\$THIDK\$WCYKRV\$VP-FG\$RPEGV\$DG	525
Qy	518	SWAKWDPYGC\$RTCGG\$VOLARRQCTNPTPANGKYK\$CEGV\$KRV\$K\$CNLEPC\$P\$SASGK	577
Db	526	ANGPWT\$W\$GDC\$RTCGG\$VSS\$SRH\$CDS\$P\$RTIG\$K\$YCL\$GER\$R\$H\$R\$SCNTDDC\$P\$G\$S--Q	583
Qy	578	SFREEQCEAFNG-----YNH\$TNRLTLAVAWVPKY\$GV\$V\$PRDK\$K\$LCIKRANGTGYFYV	630
Db	584	DFREVQCAEFDSIPFRGK\$YK\$WKTYR-----GGGVK-----AC\$SLT\$LAEG\$NFYTE	630
Qy	631	LAPKVDGTLCS\$D\$T\$V\$CVQ\$GCIKAGCDNGL\$G\$K\$K\$R\$F\$D\$K\$G\$V\$G\$G\$D\$N\$K\$S\$CK\$K\$VT\$G\$LT	690
Db	631	RAAAVVDGTPCRP\$D\$T\$D\$ICV\$G\$E\$CK\$H\$V\$G\$C\$H\$V\$G\$C\$D\$R\$V\$L\$G\$D\$R\$E\$D\$K\$R\$V\$C\$G\$D\$G\$S\$AC\$E\$T\$IE\$G\$V\$F\$S	690
Qy	691	--KPMHGYNFV\$AI\$PAGAS\$IDIR\$G\$Y\$K\$G\$LI\$G\$D\$N\$Y\$AL\$K\$N\$S\$Q\$K\$Y\$LL\$N\$G\$H\$F\$V\$V\$SAVER	748
Db	691	PASPGAGYEDV\$V\$WIP\$K\$S\$V\$H\$FIQD-----LNL\$SL\$H\$AL\$K\$G\$D\$E\$S\$LL\$E\$G\$L\$P\$G\$T\$P\$Q\$P\$H\$R	745
Qy	749	DLV\$V\$K\$S\$LL\$R\$Y\$G\$T\$G\$T\$AV\$E\$S\$Q\$AS\$R\$P\$LE\$P\$T\$V\$E\$V\$L\$S\$V\$G\$K\$M\$T\$P\$R\$V\$R\$Y\$S\$F\$Y\$LP-	801
Db	746	-LPLAGTTFQ\$R\$Q\$D\$Q\$V\$Q\$S\$LEAL\$G\$P\$IN\$AS\$LI\$V\$V\$L\$ARTEL--PAL\$Y\$R\$E\$N\$AP\$I\$ARD\$SLP	802
Qy	802	-----KEPREDK--SS\$R\$PKD	814
Db	803	PY\$W\$H\$Y\$A\$P\$W\$T\$K\$S\$A\$Q\$A\$G\$G\$Q\$V\$Q\$A\$V\$E\$C\$R\$N\$Q\$LD\$S\$A\$V\$A\$P\$H\$Y\$C\$S\$A\$H\$S\$K\$LP\$K\$Q\$R\$A\$C\$N\$T\$E\$P\$C\$P	862
Qy	815	P-----RGPSV\$L\$H\$N\$V\$L\$S\$L\$N\$Q\$VEQ-----PDD\$R\$P\$P\$A-----	841
Db	863	P\$W\$V\$V\$G\$N\$M\$S\$L\$C\$R\$S\$C\$D\$A\$G\$V\$R\$S\$T\$S\$V\$V\$C\$Q\$R\$V\$S\$A\$A\$E\$K\$A\$LD\$S\$A\$C\$P\$Q\$P\$R\$P\$V\$V\$LEA\$C\$H\$G\$T\$C\$P	922
Qy	842	-RWAG\$S\$G\$P\$C\$S\$C\$G\$S\$C\$LO\$K\$R\$A\$V\$D\$C\$R\$S\$A\$G\$O\$R\$T\$V\$P\$A--C\$D\$A\$A\$H\$R\$P\$V\$E\$T\$O\$A\$C\$G-EPC\$P--	895
Db	923	PEWATLDW\$E\$C\$T\$P\$S\$C\$G\$P\$G\$R\$H\$R\$V\$V\$C\$K\$S\$A\$Q\$O\$R\$S\$T\$LP\$P\$C\$H\$C\$L\$P\$A\$A\$K\$P\$P\$T\$W\$R\$C\$N\$R\$C\$P\$P\$A	982
Qy	896	TWEL\$S\$A\$N\$S\$P\$C\$S\$K\$G\$R\$G\$F\$Q\$R\$R\$S\$K\$C\$V\$G\$H\$G\$G\$R\$LL\$A\$R\$D\$Q\$C\$N\$L\$H\$R\$K\$P\$Q\$E\$L\$D\$F\$C	945
Db	983	RW\$T\$S\$E\$N\$G\$E\$C\$T\$Q\$C\$G\$L\$G\$Q\$Q\$O\$R\$T\$V\$R\$C\$T\$S\$H\$T\$Q--P\$S\$R\$E\$C\$T\$E\$A\$L\$R\$E\$S\$T\$M\$Q\$O\$C	1030

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:17:26 ; Search time 44 Seconds
(without alignments)
2877.002 Million cell updates/sec

Title: US-09-965-631-4
Perfect score: 5162
Sequence: I MLLGIITLAFAGTAGGSE.....DQNLHRKFOHDFCVLRPC 950

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5162	100.0	950	23	AAE22541 Human protease #2.
2	5156	99.9	950	22	AAG62299 Human metalloprotease
3	5117	99.1	952	23	AAU74751 Human protease PRT
4	4862.5	94.2	928	23	AAU72899 Human metalloprotease
5	2497.5	48.4	505	21	AAU71257 Rat metalloprotease
6	2486.5	48.2	950	20	AAU49501 Human METH1 protei
7	2486.5	48.2	950	22	AAU73549 Human ADAM-type me
8	2486.5	48.2	950	22	AAU50002 Human integrin lig
9	2486.5	48.2	967	19	AAU80285 Human integrin lig
10	2486.5	48.2	968	22	AAU50011 Protein; SHQ ID 12

11	2485.5	48.1	967	20	AAU04142 Human Tango-71 pro
12	2464.5	47.7	967	20	AAU78189 Human secreted pro
13	2448.5	47.4	950	21	AAU53899 Amino acid sequenc
14	2325	45.0	896	21	AAB21265 Mouse metalloprote
15	2274	44.1	727	20	AAU78435 Human ADAMTS-1 pro
16	2138.5	41.4	890	20	AAU49502 Human METH2 protei
17	2138.5	41.4	890	22	AAU50003 Human METH2. Homo
18	2129.5	41.3	895	22	AAU74946 Murine ADAM type me
19	2115	41.0	909	22	AAU72884 Human ADAMTS-8 am
20	1966	38.1	381	21	AAB21261 Human metalloprote
21	1965	38.1	1629	23	ABG30703 Human aggrecanase
22	1965	38.1	1629	23	AAU14448 Human ADAMTS-S1 pr
23	1965	38.1	1916	23	AAU19173 Human protease, PR
24	1965	38.1	1935	23	AAU72896 Human metalloprote
25	1957	37.9	1602	23	ABG30702 Human aggrecanase
26	1943.5	37.7	1073	21	AAB21264 Human metalloprote
27	1939.5	37.6	1934	22	AAU72301 Human ADAMTS-9 alt
28	1919	37.2	367	23	AAE22542 Human protease #3.
29	1915	37.1	837	20	AAU75425 Human aggrecan deg
30	1914	37.1	837	21	AAU99429 Human PRO1563 (UNQ
31	1914	37.1	837	22	AAU29199 Human PRO polypept
32	1914	37.1	837	22	AAU66178 Protein of the inv
33	1913	37.1	837	22	AAU78228 Human aggrecanase-
34	1913	37.1	840	21	AAB21256 Human metalloprote
35	1903	36.9	680	21	AAB21251 Human metalloprote
36	1882	36.5	1907	23	AAU71133 Human protease #12
37	1874	36.3	1505	23	AAU72897 Human metalloprote
38	1872	36.3	947	22	AAU86950 Human metalloprote
39	1854	35.9	930	22	AAU72280 Murine ADAMTS-5 am
40	1848.5	35.8	1882	22	AAU72286 Human ADAMTS-9 am
41	1847.5	35.8	930	20	AAU75426 Human aggrecan deg
42	1840	35.6	929	21	AAU41226 Human OREF ORF90
43	1798.5	34.8	874	22	AAU72287 Murine ADAMTS-9 am
44	1783	34.5	958	21	AAB21255 Human metalloprote
45	1762	34.1	870	21	AAB21252 Rat metalloprotein

ALIGNMENTS

RESULT 1
AAE22541
ID AAE22541 standard; Protein; 950 AA.

AC AAE22541;

DT 26-JUL-2002 (first entry)

DE Human protease #2.

DE Human; novel human protein; NHP; protease; biological disorder; obesity;

KW high blood pressure; arthritis; connective tissue disorder; infertility;

KW gene therapy; enzyme.

XX Homo sapiens.

XX WO200226949-A2.

XX 04-APR-2002.

XX 27-SEP-2001; 2001WO-US30350.

XX 29-SEP-2000; 2000US-236689P.

XX (LEXI-) LEXICON GENETICS INC.

XX Friddle CJ, Hilbun E;

XX WPI: 2002-372123/40.

XX N-PSDB; AAD35569.

XX Novel nucleic acid encoding a human protease, useful as a hybridization

XX probe for screening libraries and assessing gene expression patterns -

XX PS Claim 6; Page 36-38; 41pp; English.

CC The present sequence is novel human protein (NHP), human protease.

CC NHPs share structural similarity with animal proteases particularly

CC zinc metalloproteases. Sequences of the invention are useful in

CC therapeutic, diagnostic and pharmacogenomic applications. NHP

CC polynucleotides are used as hybridisation probes for screening

CC libraries and assessing gene expression patterns. They can also be

CC used for treating related biological disorders such as obesity, high

CC blood pressure, arthritis, connective tissue disorders and infertility.

CC They are also used in gene therapy.

XX XX Sequence 950 AA;

SQ Query Match 100.0%; Score 5162; DB 23; Length 950;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

DB 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTELHGLVPLQGLTGGSDLRRCFYSGDVNAEPDFAAVSLC 120

DB 61 QEDFYHLTPDAQFLAPAFSTELHGLVPLQGLTGGSDLRRCFYSGDVNAEPDFAAVSLC 120

QY 121 GGLRGAFYRGAEYVISPUNASAPAAQNSOGAHLQRRGVGGPDTSRCGVASGW 180

DB 121 GGLRGAFYRGAEYVISPUNASAPAAQNSOGAHLQRRGVGGPDTSRCGVASGW 180

QY 181 NPAILRALDPKPRRAGFSGESRRSRGRKRVSTIPRYETLVVADESVMVRFHGADLEH 240

DB 181 NPAILRALDPKPRRAGFSGESRRSRGRKRVSTIPRYETLVVADESVMVRFHGADLEH 240

QY 241 YLTLTATAARLYRHPISILNINIVVKKLLDRDSGPKVTGNAALTLENFCAMOKKLN 300

DB 241 YLTLTATAARLYRHPISILNINIVVKKLLDRDSGPKVTGNAALTLENFCAMOKKLN 300

QY 301 KVSQKHEPYWDTALLTRQDLCCATTCDTLGMADVGTMDPKRSCSVIEDDGLPSAFTTA 360

DB 301 KVSQKHEPYWDTALLTRQDLCCATTCDTLGMADVGTMDPKRSCSVIEDDGLPSAFTTA 360

QY 361 HELGHVFNMPHDNVKVEEYFVGLKLANHMSPTLIQIDRANPWSACSAALITDFLDSHG 420

DB 361 HELGHVFNMPHDNVKVEEYFVGLKLANHMSPTLIQIDRANPWSACSAALITDFLDSHG 420

QY 421 DCLLDQPSKPISLPDLPGASYTLTSGOCELAFGVGSKPCPYMOYCTKLWCIGKAKQWVC 480

DB 421 DCLLDQPSKPISLPDLPGASYTLTSGOCELAFGVGSKPCPYMOYCTKLWCIGKAKQWVC 480

QY 481 QTRHFPWADGTSCGEGKLCILKGACVERHNLKRVGDSWAKWDPYGPCSFICGGGVOLAR 540

DB 481 QTRHFPWADGTSCGEGKLCILKGACVERHNLKRVGDSWAKWDPYGPCSFICGGGVOLAR 540

QY 541 RQCTNPTPANGKYCEGVRRVVKYSCNLEPCPSASGKSPFEEQCEAFNGYHNSTNLTILA 600

DB 541 RQCTNPTPANGKYCEGVRRVVKYSCNLEPCPSASGKSPFEEQCEAFNGYHNSTNLTILA 600

QY 601 VAWPKYSGVSPDKCKLICRANGTGYYVLAKVVDGTLCSPDSTSVCGQKICRAGCD 660

DB 601 VAWPKYSGVSPDKCKLICRANGTGYYVLAKVVDGTLCSPDSTSVCGQKICRAGCD 660

QY 661 GNLGSKRFRDKCGVGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720

DB 661 GNLGSKRFRDKCGVGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720

QY 721 GDDNYLALKNSQGYLLNGHFVSAVERDLVWKSLLIRYSGTGTAVESLQASRPILLEPT 780

DB 721 GDDNYLALKNSQGYLLNGHFVSAVERDLVWKSLLIRYSGTGTAVESLQASRPILLEPT 780

QY 781 VEVLVSGKMTPPRYRYFYLPKPEDEKSSHPKDPGRGSPVHLNHSVLSLSNQVEQPDORPP 840

DB 781 VEVLVSGKMTPPRYRYFYLPKPEDEKSSHPKDPGRGSPVHLNHSVLSLSNQVEQPDORPP 840

QY 841 ARWVAGSWGPCSCSGSLQKRAVDCRGSGAGQRTVPACDAARHPVETQACGECPTWELS 900

DB 841 ARWVAGSWGPCSCSGSLQKRAVDCRGSGAGQRTVPACDAARHPVETQACGECPTWELS 900

QY 901 AWSPCSKSCGGRGFRRLSKVCGHGRLLARDQCNLHKKPOELDFCVLRPC 950

DB 901 AWSPCSKSCGGRGFRRLSKVCGHGRLLARDQCNLHKKPOELDFCVLRPC 950

RESULT 2

AAG62299

ID AAG62299 standard; protein; 950 AA.

XX AC AAG62299;

XX AC AAG62299;

DT 23-AUG-2001 (first entry)

XX DE Human metalloprotease MDTs6 protein.

XX KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;

XX KW osteopathic; antiarthritic.

XX OS Homo sapiens.

XX PN WO200134785-A1.

XX PD 17-MAY-2001.

XX PF 10-NOV-2000; 2000WO-JP07917.

XX PR 11-NOV-1999; 99JP-0321740.

PR 16-NOV-2000; 2000JP-0144020.

XX PA (YAMA) YAMANOUCHI PHARM CO LTD.

PA (KAZU-) KAZUSA DNA RES INST.

XX PI Yamaaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

DR WPI; 2001-343602/36.

DR N-PSDB; AAH41003.

XX PT Metalloprotease with aggrecanase activity for treating joint diseases

PT especially osteoarthritis

XX PS Claim 1; Page 56-60; 85pp; Japanese.

XX This invention relates to a metalloprotease with aggrecanase activity.

CC The invention includes protein and DNA sequences of the metalloprotease,

CC vectors containing the DNA, host cells transformed by the vectors, and

CC antibodies directed against the metalloprotease. The antibodies, protein

CC and DNA sequences can be used in the treatment and prevention of joint

CC diseases, particularly osteoarthritis. The treatment may result in

CC osteopathic and antiarthritic activity. The present sequence represents

CC the metalloprotease of the invention termed MDTs6.

XX SQ Sequence 950 AA;

Query Match 99.9%; Score 5156; DB 22; Length 950;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

DB 1 MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGDLIFQITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTELHGLVPLQGLTGGSDLRRCFYSGDVNAEPDFAAVSLC 120

DB 61 QEDFYHLTPDAQFLAPAFSTELHGLVPLQGLTGGSDLRRCFYSGDVNAEPDFAAVSLC 120

QY 121 GGLRGAFYRGAEYVISPUNASAPAAQNSOGAHLQRRGVGGPDTSRCGVASGW 180

Db 121 GGLRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGSPGDETSRCGVASGW 180
QY 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESANVKKFPGADLEH 240
Db 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESANVKKFPGADLEH 240
QY 241 YLLTLLATAARLYRHPSTILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKLN 300
Db 241 YLLTLLATAARLYRHPSTILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKLN 300
QY 301 KVSQKHPEYWDALFTTRQDLGGATTCDTLGMADVGTWCDPKRSCSVIEDGLPSAFTTA 360
Db 301 KVSQKHPEYWDALFTTRQDLGGATTCDTLGMADVGTWCDPKRSCSVIEDGLPSAFTTA 360
QY 361 HELGHVFNHNDNVKVCBEVEFKLRANHMSPTLIQIDRANPWSACSAALITDFLDSGHG 420
Db 361 HELGHVFNHNDNVKVCBEVEFKLRANHMSPTLIQIDRANPWSACSAALITDFLDSGHG 420
QY 421 DCLLDQPSKPIISLPEDLPASGTYTISQOCCLAFVGSKPCPYMOCYCTKLWCTGKAKGQWVC 480
Db 421 DCLLDQPSKPIISLPEDLPASGTYTISQOCCLAFVGSKPCPYMOCYCTKLWCTGKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKLCILKAGACVERHNLNKHVRDGSWAKWDYGPSCSRTCCGGVQLAR 540
Db 481 QTRHFPWADGTSCEGKLCILKAGACVERHNLNKHVRDGSWAKWDYGPSCSRTCCGGVQLAR 540
QY 541 RQCTNPTPANGKYCEGVYRKYRSCNLEPCPSSASGSKSFREEOCEAFNGYHNSTNRLTLA 600
Db 541 RQCTNPTPANGKYCEGVYRKYRSCNLEPCPSSASGSKSFREEOCEAFNGYHNSTNRLTLA 600
QY 601 VAWPKYSGVSPDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVGVGKCIKAGCD 660
Db 601 VAWPKYSGVSPDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVGVGKCIKAGCD 660
QY 661 GNLGSKKRFKCGYCGGDKNSCKKVTGLFTKPMHGYNFVVAIPAGASSIIRORGYKGLI 720
Db 661 GNLGSKKRFKCGYCGGDKNSCKKVTGLFTKPMHGYNFVVAIPAGASSIIRORGYKGLI 720
QY 721 GDDNYLALKNSQKYLNGHFVSAVERDLVWKSLLRYSCTGTAVESLCAASRPILPLT 780
Db 721 GDDNYLALKNSQKYLNGHFVSAVERDLVWKSLLRYSCTGTAVESLCAASRPILPLT 780
QY 781 VEVLVSKMTPPRVRYSEYLPKPREDKSSHDPKPCPSVLHNSVLNSQVQPDORPP 840
Db 781 VEVLVSKMTPPRVRYSEYLPKPREDKSSHDPKPCPSVLHNSVLNSQVQPDORPP 840
QY 841 ARWVAGSWGPCSCSGSLQKRAVDRCGSAGQRTVPACDAHRPVTQACGEPCTWELS 900
Db 841 ARWVAGSWGPCSCSGSLQKRAVDRCGSAGQRTVPACDAHRPVTQACGEPCTWELS 900
QY 901 AWSPCSKSGRGFORSLKCVGHGRLRLARDQCNLHRKPQELDFCVLRPD 950
Db 901 AWSPCSKSGRGFORSLKCVGHGRLRLARDQCNLHRKPQELDFCVLRPD 950

RESULT 3

ID AAU74751 standard; Protein; 952 AA.

AC AAU74751;

DT 09-APR-2002 (first entry)

DE Human protease PRTS-11 protein sequence.

XX Human; protease; PRTS; gastrointestinal; Crohn's disease; 240
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.
OS Homo sapiens.

XX WO200198468-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US19178.

XX 16-JUN-2000; 2000US-212336P.

XX 22-JUN-2000; 2000US-213955P.

XX 29-JUN-2000; 2000US-215396P.

XX 07-JUL-2000; 2000US-216821P.

XX 14-JUL-2000; 2000US-218946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;
PI Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
PI Wallia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;

XX WPI; 2002-090437/12.

XX N-PSDB; ABK12894.

XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful
in the diagnosis, treatment and prevention of gastrointestinal (e.g.
gastritis), cardiovascular (e.g. atherosclerosis) and cell
proliferative (e.g. cancer) disorders -

XX Claim 1; Page 144-146; 177pp; English.

XX The present invention relates to twenty one new human proteases,
referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and
polypeptides of the invention are useful in the diagnosis, treatment and
prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
myocardial infarction, autoimmune/inflammatory e.g. acquired
immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
proliferative e.g. cancer, developmental e.g. Duchenne and Becker
muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
epilepsy and Alzheimer's disease and reproductive e.g. infertility and
endometriosis disorders. Numerous other examples of each disorder are
given in the specification. The present protein sequence represents
the human protease PRTS-11 protein of the invention.

XX SQ Sequence 952 AA;

Query Match 99.1%; Score 5117; DB 23; Length 952;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 947; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 MLLIGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYWRGPDSDGGLIFQITAF 60

Db 1 MLLIGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYWRGPDSDGGLIFQITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRRCFYSGDVNAEPDFAAVSLC 120

Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRRCFYSGDVNAEPDFAAVSLC 120

QY 121 GGLRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGSPGSDTSSRCGVASGW 180

Db 121 GGLRGAFYRGAEYVISPLPNASAPAAQNSQGAHLQRRGVPGSPGSDTSSRCGVASGW 180

QY 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESANVKKFPGADLEH 240

Db 181 NPAILRALDPYKPRAGGSESRSSRRSGRAKREVISIPRYVETLVVADESANVKKFPGADLEH 240

QY 241 YLLTLLATAARLYRHPSTILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKLN 300

Db 241 YLLTLLATAARLYRHPSTILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKLN 300

QY 301 KVSQKHPEYWDALFTTRQDLGGATTCDTLGMADVGTWCDPKRSCSVIEDGLPSAFTTA 360

Db 301 KVSQKHPEYWDALFTTRQDLGGATTCDTLGMADVGTWCDPKRSCSVIEDGLPSAFTTA 360

Nucleic Acid
132 7473089CBI
98.8% nu Acid
100% Q.A.
SEA 20 NO 11
7473089CBI

Db	301	KVSDKHPEYWDTAITLFTRODLGATTCDTLGMADVGTMCMDPKRSCSVILEDGLPSAFTTA	360
QY	361	HELGHVFNPHDNVNVKCEEVFGKLRANHMMSPTLQIDRANPWSACSAAIITDFLDSHG	420
Db	361	HELGHVFNPHDNVNVKCEEVFGKLRANHMMSPTLQIDRANPWSACSAAIITDFLDSHG	420
QY	421	DCLLDQSPKIPSLPEDLPAGSYTLQOCELAFCVGSKPCPYMOYCTKLTWCTGKAKQWVC	480
Db	421	DCLLDQSPKIPSLPEDLPAGSYTLQOCELAFCVGSKPCPYMOYCTKLTWCTGKAKQWVC	480
QY	481	QTRHFWADGTCGEGKLCIKAGACVERHNLNHRVDGSKWAKWDPYGPCSRTCGGVQLAR	540
Db	481	QTRHFWADGTCGEGKLCIKAGACVERHNLNHRVDGSKWAKWDPYGPCSRTCGGVQLAR	540
QY	541	RQCTNPTPANGKYCEGVVRVYKSCNLEPCPSSASGKSFRFEOCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGKYCEGVVRVYKSCNLEPCPSSASGKSFRFEOCEAFNGYNHSTNRLTLA	600
QY	601	VAVVPKYSGVSPRDKCLICRANGTGYFYVLAPE-VVDGTLCPDSTSVVVOGKCIKAGC	659
Db	601	VAVVPKYSGVSPRDKCLICRANGTGYFYVLAPEVVDGTLCPDSTSVVVOGKCIKAGC	660
QY	660	DGNLGSKKRFDKCGVCGDGNKSKVVTGLTKPMHGNFVVAIPAGASSIDIRQRYKGL	719
Db	661	DGNLGSKKRFDKCGVCGDGNKSKVVTGLTKPMHGNFVVAIPAGASSIDIRQRYKGL	720
QY	720	IGDNYLALNKSQKYLNLGHFVVSVERDLVVGSLRYSGTGAVESLQASRPILPL	779
Db	721	IGDNYLALNKSQKYLNLGHFVVSVERDLVVGSLRYSGTGAVESLQASRPILPL	780
QY	780	TVEVLSVGMKTPPRVRYSFYLPKREPKSKSHPKDPR-GPSVLHNSVLSLSNOVEQPDOR	838
Db	781	TVEVLSVGMKTPPRVRYSFYLPKREPKSKSHPKDPR-GPSVLHNSVLSLSNOVEQPDOR	840
QY	839	PPARWAGSWGPCASGSGLOKRAVDCRSGAGORTVPACDAARHVPETACGECPTWE	898
Db	841	PPARWAGSWGPCASGSGLOKRAVDCRSGAGORTVPACDAARHVPETACGECPTWE	900
QY	899	LSAMSPCKSGRGFORSLKCVGHGRLARDOCNLHRKPQELDFCVLRPC	950
Db	901	LSAMSPCKSGRGFORSLKCVGHGRLARDOCNLHRKPQELDFCVLRPC	952
RESULT 4			
ID	AAU72899	(SG Pr: 168) Nucleotide	
XX	AAU72899	standard; Protein: 928 AA	
AC	AAU72899	(SG Pr: 168) Nucleotide	
XX	AAU72899	(first entry)	
DT	26-FEB-2002	Human metalloprotease partial protein sequence #11.	
DE	Human metalloprotease partial protein sequence #11.		
XX	Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;		
KW	vasotropic; antimigraine; analgesic; endocrine; neurotropic; tranquilizer;		
KW	hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;		
KW	anorectic; antiinflammatory; aspartyl protease; cysteine protease;		
KW	metalloprotease; serine protease; cancer; haematopoietic; breast;		
KW	lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;		
KW	immune-related disease; cardiovascular disease; neuronal disease;		
KW	migraine; sexual dysfunction; mood disorder; attention disorder;		
KW	cognition disorder; hypotension; hypertension; psychotic disorder;		
KW	dyskinesia; metabolic disorder; inflammatory disorder.		
OS	Homo sapiens.		
XX	WO200183782-A2.		
PN	08-NOV-2001.		
XX	04-MAY-2001; 2001WO-US14431.		
XX	04-MAY-2000; 2000US-2018799		

XX	PA	(SUGEN-) SUGEN INC.	
XX	PI	Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;	
XX	PI	Payne V;	
XX	DR	WPI; 2002-041502/05.	
XX	DR	N-PSDB; AAS97182.	
XX	PT	Novel protease polypeptide useful for screening for substances that may	
XX	PT	be used to treat, e.g., cancers, immune-related diseases,	
XX	PT	cardiovascular disease, migraine, pain, psychotic and inflammatory	
XX	PT	disorders -	
XX	PS	Claim 28; Figure 2G; 232pp; English.	
XX	CC	The invention relates to an isolated, enriched, or purified protease	
XX	CC	polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to	
XX	CC	screen for substances (S) that may modulate its activity. Administering	
XX	CC	S (which modulates protease activity in vitro) may be used to treat a	
XX	CC	disease or disorder selected from cancers (e.g., of tissues, of blood or	
XX	CC	haematopoietic origin, of the breast, colon, lung, prostate, cervical,	
XX	CC	brain, ovarian, bladder or kidney), immune-related diseases and	
XX	CC	disorders, cardiovascular disease, brain or neuronal-associated diseases	
XX	CC	(e.g., central or peripheral nervous system diseases, migraine, pain,	
XX	CC	sexual dysfunction, mood disorders, attention disorders, cognition	
XX	CC	disorders, hypotension, hypertension, psychotic disorders, neurological	
XX	CC	disorders and dyskinesias), metabolic disorders and inflammatory	
XX	CC	disorders. (I) may also be useful as a diagnostic tool for a disease or	
XX	CC	disorder such as those above. AAU72876-AAU72910 represent human	
XX	CC	protease amino acid sequences of the invention.	
XX	SQ	Sequence 928 AA;	
QY	1	MULLGTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSGDOGLFIQITAF	60
Db	26	MULLGTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSGDOGLFIQITAF	85
QY	61	QEDFYHLTPDPAQFLAPAFSTHEHGVPLQGLTGSSDLRRCFYSGDVNAEPDFAAASLC	120
Db	86	QEDFYHLTPDPAQFLAPAFSTHEHGVPLQGLTGSSDLRRCFYSGDVNAEPDFAAASLC	145
QY	121	GGIRGAFGYRGAEEYVISPPLNAPAAOANRNSOGAHLRQRRVPGPSGDPTRSCGVASGW	180
Db	146	GGIRGAFGYRGAEEYVISPPLNAPAAOANRNSOGAHLRQRRVPGPSGDPTRSCGVASGW	205
QY	181	NPAILRALDPYKPRRAGTGESRRSRSGRAKRFVSIPIRYVETLVVADSMVKFHGADLEH	240
Db	206	NPAILRALDPYKPRRAGTGESRRSRSGRAKRFVSIPIRYVETLVVADSMVKFHGADLEH	265
QY	241	YLLTLTATARLYRHPISILNIPINIVVVKVLLLRDRSDGPKVTGNAALTIRNFCANOKKLN	300
Db	266	YLLTLTATARLYRHPISILNIPINIVVVKVLLLRDRSDGPKVTGNAALTIRNFCANOKKLN	325
QY	301	KVSDKHPEYWDTAITLFTRODLGATTCDTLGMADVGTMCMDPKRSCSVILEDGLPSAFTTA	360
Db	326	KVSDKHPEYWDTAITLFTRODLGATTCDTLGMADVGTMCMDPKRSCSVILEDGLPSAFTTA	385
QY	361	HELGHVFNPHDNVNVKCEEVFGKLRANHMMSPTLQIDRANPWSACSAAIITDFLDSHG	420
Db	386	HELGHVFNPHDNVNVKCEEVFGKLRANHMMSPTLQIDRANPWSACSAAIITDFLDSHG	445
QY	421	DCLLDQSPKIPSLPEDLPAGSYTLQOCELAFCVGSKPCPYMOYCTKLTWCTGKAKQWVC	480
Db	446	DCLLDQSPKIPSLPEDLPAGSYTLQOCELAFCVGSKPCPYMOYCTKLTWCTGKAKQWVC	505
QY	481	QTRHFWADGTCGEGKLCIKAGACVERHNLNHRVDGSKWAKWDPYGPCSRTCGGVQLAR	540
Db	506	QTRHFWADGTCGEGKLCIKAGACVERHNLNHRVDGSKWAKWDPYGPCSRTCGGVQLAR	565

QY 541 ROCTNPTPANGKYCEGVYKRSNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
|||||
Db 566 ROCTNPTPANGKYCEGVYKRSNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 625
|||||
QY 601 VAWPKYSGVSPRDKKLCICRANGTYFYVLAPK-VVDGTLCSDDSTSVQVQKICAKG 659
|||||
Db 626 VAWPKYSGVSPRDKKLCICRANGTYFYVLAPKVVVDGTLCSDDSTSVQVQKICAKG 685
|||||
QY 660 DGNLGSKKRFDKCGVCGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRGKYKGL 719
|||||
Db 686 DGNLGSKKRFDKCGVCGGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRGKYKGL 745
|||||
QY 720 IGDNDYLAALKNSOGKYLKNGHFWVAVERDLVVKGSLLRYSGTGTAVERSLQASRPILLEPL 779
|||||
Db 746 IGDNDYLAALKNSOGKYLKNGHFWVAVERDLVVKGSLLRYSGTGTAVERSLQASRP----- 800
|||||
QY 780 TVEVLVSGKWTTPRVRYSYFLKPEPREDKSSHPKDPGSPVLSLSNQVQEPQDDRP 839
|||||
Db 801 -----NSVLSLSNQVQEPQDDRP 817
|||||
QY 840 PARWAGSWGPCASGSGLOKRAVDRCGSAGORTVPACDAHRPVEVTAQCEPCPTWEL 899
|||||
Db 818 PARWAGSWGPCASGSGLOKRAVDRCGSAGORTVPACDAHRPVEVTAQCEPCPTWEL 877
|||||
QY 900 SAMPSCSKSGRGFORRSKCVGHGRLRLARDQCNLHRKPQELDFCVLRQC 950
|||||
Db 878 SAMPSCSKSGRGFORRSKCVGHGRLRLARDQCNLHRKPQELDFCVLRQC 928
|||||
RESULT 5
AAB21257
ID AAB21257 standard; Protein; 505 AA.
XX
AC AAB21257;
DT
DE 23-FEB-2001 (first entry)
XX
KW Rat; ADAMTS-5; metalloproteinase; ADAM;
KW a disintegrin and metalloproteinase domain; thrombospondin domain;
KW vaccine; nontropic; neuroprotective; antiparkinsonian;
KW cerebroprotective; cytostatic; antiarthritic; immunosuppressive;
KW Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;
KW autoimmune disease; brain tumour; brain injury.
OS
XX Rattus norvegicus.
XX
PN WC2000053774-A2.
XX
PD 14-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US06237.
XX
PR 08-MAR-1999; 99US-0264585.
XX
PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
XX
XX Kelner GS, Clark M, Maki RA;
XX
XX WPI; 2000-594326/56.
XX
DR N-PSDB; AAA95827.
XX
XX Polynucleotide encoding novel members of a disintegrin,
PT metalloproteinase and thrombospondin domain protein family used to
PT prevent and treat Alzheimer's disease, cancer and autoimmune diseases -
XX
XX Claim 12; Fig 14; 129pp; English.
XX
XX The present sequence is rat metalloproteinase ADAMTS-5. The
CC ADAMTS family of proteins is closely related to the ADAM (A Disintegrin
CC and Metalloproteinase Domain) family. Members of the ADAMTS family

CC contain a thrombospondin domain in addition to the disintegrin and
CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are
CC useful for the manufacture of medicaments for treating conditions
CC associated with neuroinflammation and/or neurodegeneration, such as
CC Alzheimer's disease, Parkinson's disease and stroke. They are also
CC useful for treating conditions associated with cell proliferation, cell
CC migration, inflammation and/or angiogenesis, such as cancer, arthritis
CC and autoimmune diseases. They can be used to treat patients afflicted
CC with an invasive tumour, a brain tumour or brain injury.

XX Sequence 505 AA;

QY 89 QGLTGSSDLRRCFYSGDVNAEPDSFAAVSLCGGLRGAFYGVYVISPUNASAPAAQ 148
|||||
Db 13 QRLTGSSDLRRCFYSGDVNAEPDSFAAVSLCGGLRGAFYGVYVISPUNASAPAAQ 72
|||||
QY 149 RNSQGAHLQRRGYPGPGSGDPTSRGCVASGNWNPAILRALDPYKPRAGFGESRRRSR 208
|||||
Db 73 RNSQGAHLQRRGYPGPGSGDPTSRGCVASGNWNPAILRALDPYKPRAGFGESRRRSR 132
|||||
QY 209 RAKRFVSIPIRYVETLVVADESVMKFGADLEHYLLTLTAARLYRHPISILNPINIVVK 268
|||||
Db 133 RAKRFVSIPIRYVETLVVADESVMKFGADLEHYLLTLTAARLYRHPISILNPINIVVK 192
|||||
QY 269 VILLRDRDSGPKVTGNAALTFRNFCWOKKLNKVS DKHPEYWDTAILETRDLCGATCD 328
|||||
Db 193 VILLRDRDSGPKVTGNAALTFRNFCWOKKLNKVS DKHPEYWDTAILETRDLCGATCD 252
|||||
QY 329 TLGMADVGTMDCKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNVKNVCEEVFGKLRANH 388
|||||
Db 253 TLGMADVGTMDCKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNVKNVCEEVFGKLRANH 312
|||||
QY 389 MMSPTLIQIDRANPWSACSAAITDLDGSGHCDLDDPSKPSISLPEDLPGASTYLSQOC 448
|||||
Db 313 MMSPTLIQIDRANPWSACSAAITDLDGSGHCDLDDPSKPSISLPEDLPGASTYLSQOC 372
|||||
QY 449 ELAFGVSGRCPYMOYCTKLWCTGKAGQMVQCTRHFPWADGTSCGEGKCLKGACVERH 508
|||||
Db 373 ELAFGVSGRCPYMOYCTKLWCTGKAGQMVQCTRHFPWADGTSCGEGKCLKGACVERH 432
|||||
QY 509 NLNKHVDSWAKWDYPGCSRTCGGVQLARR--QCTNPTPANGKCYCEGVRYKRSN 566
|||||
Db 433 NPNKYRVDGPWAKWEYPGCSRTCGGGAQLARRVQATLPUP-TGGKYCEGVRYKRSN 491
|||||

RESULT 6

AA49501

ID AA49501 standard; Protein; 950 AA.

XX

AC AA49501;

XX

XX 10-JAN-2000 (first entry)

XX

XX Human METH1 protein.

XX

KW Human; METH1; METH2; anti-angiogenic; metalloproteinase thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
KW arterial-venous malformation; immune deficiency.

XX

OS Homo sapiens.

XX

XX WO9937660-A1.

PN

XX PD 29-JUL-1999. 99WO-US01313.
XX PF 22-JAN-1999; 98US-0072298.
XX PR 23-AUG-1998; 98US-0098539.
XX PR 26-AUG-1998; 98US-0098539.
XX PA (TRUE/) IRUELA-ARISPE L.
XX PA (HAST/) HASTINGS G A.
XX PA (RUBE/) RUBEN S M.
XX PI IrueLa-Arispe L, Hastings GA, Ruben SM;
XX WI: 1999-590684/50.
XX DR N-PSDB; AA232000.
XX PT New isolated metalloprotease thrombospondin polypeptides, useful for
XX PT treating hyperproliferative disorders, cancers or autoimmune disorders
XX FS Claim 10; Fig 1; 457pp; English.
XX AA232000 and AA232001 encode, and AA49501 and AA49502 represent, human
XX metalloprotease thrombospondin (METH) proteins METH1 and METH2
XX respectively. METH1 and METH2 have been found to be potent inhibitors of
XX angiogenesis both in vitro and in vivo. They can be used for treating
XX cancer and other disorders related to angiogenesis including abnormal
XX wound healing, inflammation, rheumatoid arthritis, psoriasis,
XX endometrial bleeding disorders, diabetic retinopathy, some forms of
XX macula degeneration, haemangiomas, and arterial-venous malformations.
XX They may be useful in treating deficiencies or disorders of the immune
XX system, by activating or inhibiting the proliferation, differentiation,
XX or mobilization (chemotaxis) of immune cells. The etiology of these
XX immune deficiencies or disorders may be genetic, somatic, such as
XX cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
XX toxins), or infectious. They can also be used to treat inflammatory
XX conditions, both chronic and acute conditions. The products can also be
XX used for detection and diagnosis. AA232002 to AA232080, and AA49503 to
XX AA49511 represent sequences given in the exemplification of the present
XX invention.
XX SQ Sequence 950 AA;
Query Match 48.2%; Score 2486.5; DB 20; Length 950;
Best Local Similarity 49.0%; Pred. No. 7.7e-184;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;
QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPDGQGLIFQTAF 60
DB 19 LLLLAALLAVSDALGRPSEDELVVP-ELE-----RAP---GHGTTLRRLHAF 64
QY 61 QEDFVHLTPDAQLTAPAFSTHGLG-----VPLQGLTGGSSDLRRRCFYSGDVNAEPDSF 114
DB 65 DQDLLELRPDSSFLAPGFTLQNVGRKSGSETPLD-----ETDLAHFYSTVNGDSSA 119
QY 115 AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQCA---HLI0--RRGVPGG 165
DB 120 AALSCEGVGAFYLLGEAYFTIQLPAAASERLATAAPCEKPPAPLQFHLIRNRQGDVGG 179
QY 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRAGFESRRRSR 207
DB 180 TCGVVDDEFRPTGKAETEDEGTEGEPQWS-----PODPALQGVGP-TGTGS 230
QY 208 GRAKRFVSTIPRVETVLVADSMVFGADLEHLYLLTLLATAARLYRHPSTLNINIVV 267
DB 231 IRKKRFVSHRYVETMLVADQSMAEFGSLGKHYLLTFLSVAARLYKHPSTIRNSVSLVV 290
QY 268 KYLLLRDRDSGPKVTGNAAALTIRNFCAMOKLNKVKHPEYWDFTAILFTRODLGGATTC 327
DB 291 KILVTHDEQKGPVTSNAALTIRNFCNMQKHNPPSDRDAEHDYDFTAILFTRODLGSGTQC 350
QY 328 DTGLMADVCTMCDPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVAVCEVEFGKLRAN 387

DB 351 DTGLMADVCTVCDPRKSCSVIEDDGLQAAFTTAHELGHVFNPHDNNVAVCEVEFGKLRAN 410
QY 388 HMMSTPLTQIDRANPWSACSAAITLDFLDGSHGDCLLDOPSKPISLPEDLPGASYTLISQQ 447
DB 411 HMMASMLNLDHISQFWSPCSAIMITSLFDNGHGECLMDKPNPQLPGLDLPCTSYDANRQ 470
QY 448 CELAFGVGSKPCP-YMOYCTKLWCTGKAKAGQMVQCTRHFPWADGTSCGEGKLCIKGACVE 506
DB 471 COFTFGEDSKHCPDAASTCSTLWCTGTSGGVLCVQTKHFPWADGTSCGEGKWCINGKCVN 530
QY 507 RHNLNKH---RVDSGWAQWDPYPCSTRTCGGVQLARQCTNPTPANGKKYCEGVRYKYR 563
DB 531 KTD-RKHFDTPFHGSMGMPGDCSRTCCGGVQVQTMRECDNPVPKNGKYCEGKRVRYR 589
QY 564 SCNLEPCPSSASGKSFREEOCAFNGYNHSTNRLTLAVAWVPKYSGVSPDRCKLCICRAN 623
DB 590 SCNLEDCFDN-NGKTFREOCEAHNEFSKASGSPAVEMIPKYAGVSPKDRCKLICCAK 648
QY 624 GTGYFVILAPKVVDGTLCSPDSTSVGVGKCIKAGCDGNLGSKKRFPDKCGVCGGDNKSK 683
DB 649 GIGYFVILQPKVVDGTPCSPDSTSVGVGQCVKAGCDRIIDSKKFKDKCGVCGGNGSTCK 708
QY 684 KYTGFTKPMHGYNFWAIPAGASSIDIRQGYKGLIGDNYLALKNQSGKYLNGHFV 743
DB 709 KISGSVTSAPGYHDIITPTGATNIEVKQRNQRNNGSFLAIAKAADGTYLLNGDYTL 768
QY 744 SAVERDLVVKSLRYSGTGTAVESLQASRPILPTVEVLSVGRMTPPRVYSFYLPKE 803
DB 769 STLEQDINMYKGVLYRSGSSAALIRSFSPLEPLTIQVLTGVNLRPKIKYTVFVKKK 828
QY 804 PREDKSSHPKDPGRPSVLNSVLSNQVEQDPPRPAWVAGSWGSPCSASCGSLQKRA 863
DB 829 ---KES-----FNAIPTF-----AWTTEWGECSKCELGWQRRL 861
QY 864 VDCRGAGORTYACDAAH--RPVETQAGE-PCPTWELSAWSPGSKSGRGFQRRSLKC 920
DB 862 VECRDINGO---PASECAKEVPAPSTRPCADHPCPQWOLGEWSSCKTGKGYKKRSLKC 918
QY 921 VHGGRLLARDOCNLHRKPQE-LDFCVLRPC 950
DB 919 LSHDGGVLSHESCDPLKPKHFDICTMAEC 949
RESULT 7
AA232000
ID AAB73549 standard; Protein; 950 AA.
XX AAB73549;
AC AAB73549;
DT 07-AUG-2001 (first entry)
XX Human ADAM-type metalloprotease MDT54, SEQ ID NO:4.
DE Human: MDT54; ADAM-type metalloprotease; drug screening;
KW A Disintegrin And Metalloprotease; cancer; arthritis.
XX Homo sapiens.
XX JP2001017183-A.
XX 23-JAN-2001.
XX 09-JUL-1999; 99JP-0196584.
XX 09-JUL-1999; 99JP-0196584.
XX (YAMA) YAMANOUCHI PHARM CO LTD.
XX WPI; 2001-275950/29.
XX N-PSDB; AA20224.
XX A new metal protease and its preparation for use as an anti-cancer and
PT

PT anti-arthritis therapeutic -

PS Claim 1; Page 12-14; 22pp; Japanese.

CC The invention relates to the novel human ADAM (A Disintegrin And Metalloprotease)-type metalloproteases MDTs4 (AAB73549) and MDTs5 (AAB73550). The metalloproteases can be used for the treatment of cancers and arthritis. The invention also relates to the genes encoding MDTs4 and MDTs5, vectors and host cells containing the MDTs4 or MDTs5 genes, the recombinant production of MDTs4 and MDTs5, and antibody specific for MDTs4 or MDTs5, and methods of screening for compounds which modulate the activity of MDTs4 and/or MDTs5. The present sequence represents human MDTs4.

XX Sequence 950 AA;

Query Match 48.28; Score 2486.5; D8 22; Length 950;
Best Local Similarity 49.08; Pred. No. 7.7e-184;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGQGLIFQITAF 60
DB 19 LLLAAALAVSDALGRSEDEELVVP-ELE-----RAP---GHGTRLRKHAF 64
QY 61 QEDFVLHLTPDAQFLAPAFSTHELG-----VPLQGLTGGSSDLRCFYSCDVNAEPDSF 114
DB 65 DQQLDLRLPSSFLAPGFTIQNGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 119
QY 115 AAVSLGGLRGAFGYRGAEYISPLPNAS---APAAQRNSOGA---HLLQ---RRGVPGG 165
DB 120 AALSCEGVRGAFYLLGEAYETQPLPAASERLATAAPGEPAPAPLQPHLLKRNQGDVG 179
QY 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPKPRAPGFSRERRS 207
DB 180 TCGVVDDEPRFTGKAETDEDEGEDEGPQWS-----PODPALQGVGP-TGTGS 230
QY 208 GRAKRFVSTPRVETLVVVADESMVKFPGADLEHYLLTLLATAAKLYRHPSLNPNIVVV 267
DB 231 IRKKREVSRRHYVETMLVADQSMAEFGSLKHYLLTFLSVAARLYKHPSIRNSVSLVV 290
QY 268 KVLRLDRDSGPKVTGNAAFLRNFCAWQKLNKVSQKHPEYWDITALLFTQDLGCATTC 327
DB 291 KILVTHDFQKPEVTSNAALTNRNFCNWKQHPNSDRAPEHYDTAILFTKQDLGSGQC 350
QY 328 DTGLMADVTGMDPKRRCSVTEDDGLPSAFTAHLEGHVFNMPHDNVKVCLEVEFGKLRAN 387
DB 351 DTGLMADVTGMDPKRRCSVTEDDGLQAAFTAHLEGHVFNMPHDNVKVCLEVEFGKLRAN 410
QY 388 HMMSTPLQIDRANPWSACSAITDFLDSHGDCCLLDQPSKPSLPELDPGASYTLISQQ 447
DB 411 HMMASMLNLDHSQSPWSPSCSAYMITSFLLDNGHGECLMDKQNPQLQPLGDLPGTSYDANRQ 470
QY 448 CELAFGVGSKPCP-YMQVCTKLWCTGKAKGQMVQCTRHFPAWAGTSCGEGKLGKACVE 506
DB 471 CQFTGEDSKHCPDRAASTCSILWCTGTSGLVLCVQTHFPAWAGTSCGEGKLGKACVE 530
QY 507 RHNLNKH---RVDGSAWAKWDYPGSCRTCCGGVQLARRQCFTNPANGAYCYCEGRVKYR 563
DB 531 KTD-RKHFTDPHGSWMGMPWGDSCRTCCGGVQYTWRECDNPVPKNGKAYCYCEGRVKYR 589
QY 564 SCNLEPCSSASGKSFREOCEAFNGYNHSTNRLTLAVAVPKYSGVSPRDKCKLICRAN 623
DB 590 SCNLEDCPN-NKGTFRQECCAHNEFSKASFGSGPAVEWIPKYAGVSPKDRKLCICQAK 648
QY 624 GTGYFVLAPKVVDTGLCPDSTSVYVQKICAGCDGNLGSKKRFPDQGVCGGDNKSK 683
DB 649 GIGYFVLQPKVVDGTPCSPDSTSVYVQKICAGCDRIIDSKKKFKDQGVCGGNGSTCK 708
QY 684 KYTGFTKPMHGYNFWAIPAGASSIDIRQYKGLIGLDNNYLAKNSQKYLNGHFVV 743
DB 709 KISGVSITAKPGYHDIITPTGATNIEYKQRNQRNGSRNGSFLAKAADETIYILNGDYTL 768
QY 744 SAVERDLVVKGSLLRYSGTGTAVESLQASRPILPLTVEVLSVGKMTPPKRVYSFLPKE 803

DB 769 STLEQDIMYKGVLRYSGSSAALERIRSFPLKEPTLIQVLTGVNALRPKIKYTFVKKK 828
QY 804 PREDKSSHPKDPGSPSVLHNSLSNQVQPDPPRPARWAGSWGPCASCSGSLQKRA 863
DB 829 ----KES-----FNAITFS-----AWIEWGECSKSCSLGWQRL 861
QY 864 VDCRSAGORTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSGCGRGFORSLKC 920
DB 862 VECRDINGQ---PASECAKEVKPASTRPCADHPQWOLGSEWSCSKTCGKGYKKSLLAC 918
QY 921 VHGGRLLARDQCNLHRKPQB-LDFCVLRPC 950
DB 919 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 949

RESULT 8

AAB50002

ID AAB50002 standard; Protein: 950 AA.

XX AC AAB50002;

XX DT 19-MAR-2001 (first entry)

XX DE Human METH1.

XX Human; METH1: metalloprotease; thrombospondin; angiogenesis inhibition;
cancer therapy; benign tumour; ocular angiogenic disease;
rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Webber syndrome;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; fibromuscular dysplasia; wound granulation;
KW Crohn's disease; atherosclerosis; birth control.

XX OS Homo sapiens.

XX PN WO200071577-A1.

XX PD 30-NOV-2000.

XX PF 25-MAY-2000; 2000WO-US14462.

XX PR 25-MAY-1999; 99US-0318208.

XX PR 20-JUL-1999; 99US-0144882.

XX PR 10-AUG-1999; 99US-0147823.

XX PR 13-AUG-1999; 99US-0373658.

XX PR 22-DEC-1999; 99US-0171503.

XX PR 22-FEB-2000; 2000US-0183792.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PA (IRUE/) IRUELA-ARISPE L.

XX PA (HAST-) HASTINGS G A.

XX PA (RUBE/) RUBEN S M.

XX PA (JONA/) JONAK Z L.

XX PA (TRUL/) TRULLI S H.

XX PA (FORN/) FORNWALD J A.

XX PA (TERR/) TERRITT J A.

XX PI IrueLa-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;

XX PI Fornwald JA, Terrett JA;

XX DR WPI: 2001-025136/03.

XX DR N-PSDB; AAC90057.

XX METH1 and METH2 polynucleotides and encoded polypeptides, used to
inhibit angiogenesis in the treatment of disorders such as cancer,
rheumatoid arthritis and psoriasis -

PS Claim 15; Fig 1; 768pp; English.

XX The present sequence is human METH1 (ME for metalloprotease and TH for thrombospondin). METH1 can be used for inhibiting angiogenesis in an individual, and for treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis, vasculogenesis, granulations, hypertrophic scars, nonunion fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, fibromuscular dysplasia, wound granulation, Crohn's disease or atherosclerosis. METH1 can also be used in birth control. METH1 can also be used in diagnostic methods for the prognosis of cancer.

XX Sequence 950 AA;

Query Match 48.2%; Score 2486.5; DB 22; Length 950;

Best Local Similarity 49.0%; Pred. No. 7.7e-184;

Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

Qy 1 MLLGILTLAFAGTAGGSPEREVVVPIRLDPDINGRRYWRGPEDSJDOGLIFQITAF 60
Db LLLIAAALLAVSALGRSPSEDEELVVP-ELE-----RAP---GHGTRLRLEHAP 64
Qy 61 QEDFLHLTPDAOPLAFSTFHLG-----VPLQGLTGGSDLRRCFYSGDVNAEPDSF 114
Db DQDLLELRDPSSFLAGFTLQVGRKSGSETPLP-----ETDLACFFSGTVNGDPSSA 119
Qy 115 AAVSLCGLRGAFGYRGAEEYVISPPLNAPN---APAAQNSQGA---HLQ---RRGVPG 165
Db AALSICEGVGAFYLLGEAYFIOPLPAASERLATAPEKPPAPLOPHLLRRNRQDVG 179
Qy 166 PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRACFGESRRRS 207
Db TCGVVDDEPRPTGAEDEDECTEGDEGPQMS-----PODPALQVGGP-TGTGS 230
Qy 208 GRARFRFIPRYVETLVVADESVMKFGADLEHYLLTLLATAARLYRHPILNIPINIVV 267
Db IRKRFRVSHRYVETLVVADESVMKFGADLEHYLLTLLATAARLYRHPILNIPINIVV 290
Qy 268 KVLRLDRSDGPKVTGNAALTLRNFCAWOKKLNKVSDBEYWDITAILFTQDLCAATC 327
Db KILVTHDEKQPEVTSNAALTLRNFCAWOKKLNKVSDBEYWDITAILFTQDLCAATC 350
Qy 328 DTLMADVGTWCDPKRSCVTEDDGLSPAFTHAELGHVFNMPHDNVKVEEVFGKLRAN 387
Db DTLMADVGTWCDPKRSCVTEDDGLSPAFTHAELGHVFNMPHDNVKVEEVFGKLRAN 410
Qy 388 HMSPTLIQIDRANPWSACSAIITDLDGSGDCLLDQPSKPISLPEIIPGASYTLISQ 447
Db HMASMLNLDHSPWSPSCSAIITDLDGSGDCLLDQPSKPISLPEIIPGASYTLISQ 470
Qy 448 CELAFGVSKPCP-YMYCTKLTCTGAKGQWVOTRHPWADGTSGEKLCLKGACVE 506
Db COFTFGEDSKHCPDAACTSLTWTCTGGVVLVCOTHFHWADGTSGEKLCLKGACVE 530
Qy 507 RHNLNKH---RVDSGKAWDPYPCSRCTGGGVQVLAARQCTNPTPANGRYCEGVRYKYR 563
Db KTD-RKHEDTFHSGWGMWGPWDCSRCTGGGVQVLAARQCTNPTPANGRYCEGVRYKYR 589
Qy 564 SCNLEPCSSASGSKFRQEQCEANGYNHSTNRLTLAVAVPKYSGVSPDKCKLICRAN 623
Db SCNLEPCSSASGSKFRQEQCEANGYNHSTNRLTLAVAVPKYSGVSPDKCKLICRAN 648
Qy 624 GTGYEYVLAPKVDGTCLSPDTSVCVGKCIKAGCDNGLSGKKRDKCGVCGGDNKSK 683
Db GTGYEYVLAPKVDGTCLSPDTSVCVGKCIKAGCDNGLSGKKRDKCGVCGGDNKSK 708
Qy 684 KVTGLTFPMHGYNFVAIPAAGASSIDIRQRYKGLIGDDNYLALKNQSKYLINGHFV 743
Db KVTGLTFPMHGYNFVAIPAAGASSIDIRQRYKGLIGDDNYLALKNQSKYLINGHFV 768

Qy 744 SAVERDLVVKSLRLYSGTGTAVESLQASRILEPLVEVLVSVCKMTPPRVRSFYLPKE 803
Db STLEQDIMYKGVLRYSGSSAALIRISFPLKEPLTIQVLTIVGNALRPKIKTYFVKK 828
Qy 804 PREKSSHPKDPGSPVLRHSVLSLSNQVEQDDPPARVWAGSWGPCSCSGSLQKRA 863
Db ---KES-----FNAIPTFS-----AWIEWGBCSKSCELGWORRL 861
Qy 864 VDCRSAGORTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSCGSGFQRRSLK 920
Db VECDINGQ---PASECAKEVKPASTRPCADHPCPQOLGESSCSTCGKGYKKRSLK 918
Qy 921 VGHGRLIARDQCNLHRKQPS-LDFCVLRPC 950
Db LSHDGVLSHESCDPLKKPKHFIDFCTMAEC 949

RESULT 9

AAW80285

ID AAW80285 standard; Protein; 967 AA.

XX AAW80285;

DT 19-JAN-1999 (first entry)

XX Human integrin ligand polypeptide ITGL-TSP.

XX ITGL-TSP; integrin ligand; angiogenic disease; cancer; atherosclerosis;
KW chronic inflammatory disorder; rheumatoid arthritis; tissue remodeling;
KW macular degeneration; diabetic retinopathy; Alzheimer's disease; human;
XX restenosis.

OS Homo sapiens.

PN EP874050-A2.

PD 28-OCT-1998.

PF 27-JAN-1998; 98EP-0300575.

PR 24-APR-1997; 97US-0845496.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (SMK) SMITHKLINE BEECHAM CORP.

PA (SMK) SMITHKLINE BEECHAM PLC.

XX Fronwald JA, Hastings GA, Jonak ZL, Terrett JA;

PI Trulli SH;

XX WPI: 1998-544643/47.

DR N-PSDB; AAV66508.

XX DNA encoding integrin ligand polypeptide ITGL-TSP - used to treat
XX angiogenic diseases, restenosis, Alzheimer's disease and in tissue
XX remodeling

XX Claim 11; Pages 6-9; 24pp; English.

XX This represents a human integrin ligand polypeptide ITGL-TSP. ITGL-TSP
XX polypeptides can be used in the treatment of angiogenic diseases such as
XX cancer, cancer metastasis, chronic inflammatory disorders, rheumatoid
XX arthritis, atherosclerosis, macular degeneration or diabetic retinopathy,
XX restenosis, Alzheimer's disease and tissue remodeling. They can be used
XX to treat a subject in need of enhanced activity or expression of the
XX ITGL-TSP polypeptide.

XX Sequence 967 AA;

Query Match 48.2%; Score 2486.5; DB 19; Length 967;

Best Local Similarity 49.0%; Pred. No. 7.9e-184;

Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

Qy 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPEDSJDOGLIFQITAF 60

Db 36 LLLLLAALLAVSDALGRSEDEELVVP-ELE-----RAP-----GCTTTLRLHAP 81
QY 61 QEDFVHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSSDLRRFYSGDVNAEPDSF 114
Db 82 DQDLLELRPOSSFLAPGFTLQNGRKSGSETPLP-----EIDLHAHCYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFYGRGAEYVISPPLNAS-----APAAORNSOGA-----HLIQ--RRGVPGG 165
Db 137 AALSCEGVRGAFYLLGAYFIQPLPRASERLATAAAGEXPAPLQPHLLLRNRQGVGG 196
QY 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRRAGFEGESRRRS 207
Db 197 TCGVVDDPRPTGKAETDEDEGTEGEDEGPQWS-----PQDPAQLQVQGP-TGTGS 247
QY 208 GRKRFRVSIPIRYVETLVVVADESVMKFGADLEHVLTLTLLATAARLYRHPISILNPINTVVV 267
Db 248 LRKRFRVSSHYRYVETLMVADDSMAEFHGLGKHYLLTLFVVAARLYRHPISILNPINTVVV 307
QY 268 KVLRLDRDSGPKVTGNAALTLRNFCAWQKLNKVSDBKHPYWDTAILFTKQDLGCGATTC 327
Db 308 KILVHIDRQKPEVTSNAALTLRNFCAWQKLNKVSDBKHPYWDTAILFTKQDLGCGATTC 367
QY 328 DTGLGADVGTMCDPKRSCTVIEDGLPSAFTTAHELGHVFNMPHDNVKVCVEEVFGKLRAN 387
Db 368 DTGLGADVGTMCDPKRSCTVIEDGLPSAFTTAHELGHVFNMPHDNVKVCVEEVFGKLRAN 427
QY 388 HMSFTLIQIDRANPWSACSAAITLDFDLSGSHGCDLQPSKPTSLPDPGPGASYTLTQQ 447
Db 428 HMASMLNLDHSPQSPCSAYMTSFLDNGHGECLMDKFNQPLQPLDGPSTSYDANRQ 487
QY 448 CELAFGVSGRCP- YMOYTKLWCTGKAKQMVQCTRHFPWADGTCGEGKCLKLGACVE 506
Db 488 COFTFGEDSKHCPDAASTCTLWGTSGGLVLCQTKHFPWADGTCGEGKCLKLGACVE 547
QY 507 RHNLNKH---RVGDSWAKWDPYPCSTCGGVLQARQCTNTPANGKAYCGVRVYR 563
Db 548 KTD-RKFDTPFGHSGWGMGPMGDCSRTCGGVOYVTRMCDNPNVNGKGYCEGKVRVYR 606
QY 564 SCNLEPCSSAGSKSFRFEEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCLICRAN 623
Db 607 SCNLEPCPN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKAYGSPADRCCLICQAK 665
QY 624 GTGVFYLAKVVDGTLCSDPSTSVVQVQKCIKAGDGNLGSKKRFDKCVCGGDNKSK 683
Db 666 GIGYFFVLQPKVVDGTLCSDPSTSVVQVQKCIKAGDGNLGSKKRFDKCVCGGDNKSK 725
QY 684 KVTGLFTKPMHGNFVVAIPAGASSIDIRQYKGLIGDNNYLAKNSQKYLINGHFVV 743
Db 726 KISGVTSAPGYHDIITPTGATNEVKORNGRNNNGSFLAIKAADGTYILNGDYL 785
QY 744 SAVERDLVWGLRLYSGTGTAVESLOASRPILPTVEVLSYGMKTPPFRYSFYLPKE 803
Db 786 STLEQDIMYGVVLYRYSGSAALERTSRFPLKEPLTQVLTGVNALRPKIKYTYFVKK 845
QY 804 PREDKSHHPDRGPRSVLHNSVLSLNOVQPDPRPARVAGSWGPCSLGSGGLQKRA 863
Db 846 ---KES-----FNAITFS-----AWVIEEWECSSCELGWQRL 878
QY 864 VDCRGSAGQRTVPACDAAH---RVETQACGE-PCPTWELSAWSPCKSGRSGFORSLAC 920
Db 879 VECRDINGQ---PASECAKEVKPASTRCPADHPCPQWQGEWSSCKTGKGYKRSKLC 935
QY 921 VGHGGRLLARDQCNLHRKQGE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKPKHFDICTMAEC 966

RESULT 10

AAB50011

ID AAB50011 standard; Protein; 968 AA.

XX

AC AAB50011;

XX DT 19-MAR-2001 (first entry)
XX DE Protein; SEQ ID 125.
XX KW Human; METH1; metalloprotease; thrombospondin; angiogenesis inhibition;
KW cancer therapy; benign tumour; ocular angiogenic disease;
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Webber syndrome;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; fibromuscular dysplasia; wound granulation;
KW Crohn's disease; atherosclerosis; birth control.
XX OS Homo sapiens.
XX PN WO200071577-A1.
XX PD 30-NOV-2000.
XX PF 25-MAY-2000; 2000WO-US14462.
XX PR 25-MAY-1999; 99US-0318208.
PR 20-JUL-1999; 99US-0144882.
PR 10-AUG-1999; 99US-0147823.
PR 13-AUG-1999; 99US-0373658.
PR 22-DEC-1999; 99US-0171503.
PR 22-FEB-2000; 2000US-0183792.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (SMK) SMITHKLINE BEECHAM CORP.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA (IRUE/) IRUELA-ARISPE L.
PA (RUBE/) HASTINGS G A.
PA (JONA/) JONAK Z L.
PA (TRUL/) TRULLI S H.
PA (FORN/) FORNWALD J A.
PA (TERR/) TERRETT J A.
XX IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
PI Fornwald JA, Terrett JA;
XX WPI, 2001-025136/03.
XX METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT inhibit angiogenesis in the treatment of disorders such as cancer,
PT rheumatoid arthritis and psoriasis -
XX Claim 15; Pages 759-763; 768pp; English.
XX The present invention relates to human METH1 and METH2 (ME for
CC metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC METH can be used for inhibiting angiogenesis in an individual, and for
CC treating cancer, benign tumours, an ocular angiogenic disease,
CC rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
CC fibromuscular dysplasia, wound granulation, Crohn's disease or
CC atherosclerosis. METH can also be used in birth control. METH can also
CC be used in diagnostic methods for the prognosis of cancer. The present
CC sequence is a protein isolated in the present invention.
XX SQ Sequence 968 AA;

Query Match 48.2%; Score 2486.5; DB 22; Length 968;
Best Local Similarity 49.0%; Pred. No. 7.9e-184;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

Db 308 KILVTHDQKGPVTSNAALTRNFCNWKQKOHNPSPDRDAEHYDTAILFTVHODLCGSGTC 367
QY 328 DTGLMADVTWCDPKRKSIVTEDGLPSAFTTAHELGHVFNMPHDNVKVLVEVEGKLRAN 387
Db 368 DTGLMADVTWCDPKRKSIVTEDGLPSAFTTAHELGHVFNMPHDNAKQASLNGVQDS 427
QY 388 HMMSTPLQIDRANPWSACSAIIITDILDSHGDCILDDQSPKISLPEDIPGASYTILSQO 447
Db 428 HMMASMLSNLDHSPWSPSCSAYMITSFLDNGHCECLMDKPNPQLPGLDLPSTSYDANRQ 487
QY 448 CELAFGVGSKPCP-YMOYCTKLWCTGAKAGOMVQCTRHFWADGTSCEGSKLKLKACVE 506
Db 488 COFTGEDSKPCPDAASTCTSLWCTGSGGVLCVQCTHFWADGTSCEGSKWCKNGKCN 547
QY 507 RHNLNKH---RVDSWAKWDYPCSRCTCGGGVQLARRQCTNPANGKAYCEGVRYKVR 563
Db 548 KTD-RKHFDTPFHSGSWGMPGDCSRCTCGGVQYTHRECDNVPKNGGYCEGKRVYR 606
QY 564 SCNLEPCSSASGKSFREEQCEAFNGYNHSTNRUTLAVAVVPKYGVSPKDKCKLICRAN 623
Db 607 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKAVGSPKDRCKLICQAK 665
QY 624 GTGYFYVLAPKVVDTCLSPDSTVCGVQKCIKAGCDENLGSKKRDKQCVCGGDNKSKC 683
Db 666 GIGYFFVLQPKVDTGTPCSPDSTVCGVQGVKACCDRIIDSKKKFKDKCVCGGNGSTCK 725
QY 684 KVTGLTKPMHGYNVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQSKYLLNGHFVV 743
Db 726 KISGVSISAKPGYHDIITPTGATNIEVKQNRQNSGFLAIRKADITYILNGDYTL 785
QY 744 SAVERDLVVKGLSLRYSGTAVESLQASRPLEPLTVESLVCKMTTPPVRYSFYLPKE 803
Db 786 STLEQDLIMYGWLVRYSGSAAALIRSFSPLEPLTQVLTGVALRPKIKYTFYVKK 845
QY 804 PREDKSHPKDPRGSPVLSNLSNOVEQPDPRPARVAGSWGPCSCSAGSLQKRA 863
Db 846 ---KES-----FNAITFS-----AWIEWEGCSKSCELGWORL 878
QY 864 VDCRGSGQRTVPACDAH--RPVETQACE-PCPTWELSAWSPCSKSGRGFORSLKC 920
Db 879 VECRDINGQ---PASECAKEVKPASTRPCADHPCPQWOLGEWSSCKTCKGKYKSLKC 935
QY 921 VGHGRLRLARDQCNLHKKPQE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKKPKHFDICTMAEC 966

RESULT 12
AAW78189
XX AAW78189 standard; Protein; 967 AA.
XX AAW78189;
XX
13-APR-1999 (first entry)
XX Human secreted protein encoded by gene 64 clone H0UCQ17.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 40
FT /label= unknown
FT Misc-difference 45
FT /label= unknown

FT Misc-difference 169 /label= unknown
FT Misc-difference 293 /label= unknown
FT Misc-difference 297 /label= unknown
FT Misc-difference 557 /label= unknown
XX
PN W09856804-A1.
XX
XX 17-DEC-1998.
XX
PF 11-JUN-1998; 98WO-US12125.
XX 02-OCT-1997; 97US-0061060.
PR 13-JUN-1997; 97US-0049547.
PR 13-JUN-1997; 97US-0049548.
PR 13-JUN-1997; 97US-0049549.
PR 13-JUN-1997; 97US-0049550.
PR 13-JUN-1997; 97US-0049606.
PR 13-JUN-1997; 97US-0049607.
PR 13-JUN-1997; 97US-0049608.
PR 13-JUN-1997; 97US-0049609.
PR 13-JUN-1997; 97US-0049610.
PR 13-JUN-1997; 97US-0049611.
PR 13-JUN-1997; 97US-0050566.
PR 13-JUN-1997; 97US-0050901.
PR 13-JUN-1997; 97US-0052989.
PR 08-JUL-1997; 97US-0051919.
PR 18-AUG-1997; 97US-0055984.
PR 12-SEP-1997; 97US-0058665.
PR 12-SEP-1997; 97US-0058668.
PR 12-SEP-1997; 97US-0058669.
PR 12-SEP-1997; 97US-0058750.
PR 12-SEP-1997; 97US-0058971.
PR 12-SEP-1997; 97US-0058972.
PR 12-SEP-1997; 97US-0058975.
PR 02-OCT-1997; 97US-0060834.
PR 02-OCT-1997; 97US-0060841.
PR 02-OCT-1997; 97US-0060844.
PR 02-OCT-1997; 97US-0060865.
PR 02-OCT-1997; 97US-0061059.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;
PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
PI Yu GL;
XX
XX WPI: 1999-080881/07.
XX N-PSDB: AAX04374.
XX
XX New isolated human genes and the secreted polypeptides they encode -
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders
XX
XX Claim 11; Page 297-300; 380pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin FC
XX portion (e.g. AAX04302) for increasing the stability of the fused
XX protein as compared to the human protein only.
XX The invention relates to 86 novel genes and their fragments (nucleic
XX acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein or gene therapy. Also, pathological
XX conditions can be diagnosed by determining the amount of the new
XX polypeptides in a sample or by determining the presence of mutations in
XX the new polynucleotides. Specific uses are described for each of the 86
XX polynucleotides, based on which tissues they are most highly expressed in
XX (see AAX04311 for described uses).

Db 122 DPGSAALSLCVRGAEGYVLOGEEFFIQAPGVATERLAPAVPEESSARQFHILRR- 180
 QY 162 VPGGSGDPTSCGIVASGWNPAIRLALDPYKPRRAGFESRR- 206
 Db 181 -----ARGSGKACGVMD-----DETLPSTDSRSESQNTNQPVPFDPDTPQDAGK 226
 QY 207 -----SGRAKRFVSPRYVETLVVADESVMKFGADLEHYLLTLATARLVHPHSILNPI 262
 Db 227 SGPGSIRKRFVSSPRYVETLVVADQSMADFGSLGKHYLLTLFVSAARFKHFSIRNSI 286
 QY 263 NIVVVKVLLLRDRSGPKVTGNAALTNRNFCAMQKLNKVDKHEPYMDIAILFTRQDLC 322
 Db 287 SLVVVKILVIYEEQKGPVTSNAALTNRNFCWQKNSPSDRDPEDHYDTAILFTRQDLC 346
 QY 323 GATTCDTLGMADVGTMDPKSCSVIEDDGLPSAFTTAHELGHVFNPHNVKVCVEVFG 382
 Db 347 GSHTCDTLGMADVGTCDPSRSCSVIEDDG-0AAFTTAHELGHVFNPHNPH:DAKHCASLNG 405
 QY 383 KLRANHMSPTLIQIDRANPWSACSAIITDPLDSHGDCDCLLDQPSKPISELPEDLPCASY 442
 Db 406 VSGSHLMASMLSLDHSQWSPSCSAYWVTSFLDNGHGECMDKPNQPIKIPSLDPLGTL 465
 QY 443 TLSQOCELAFVGGKPCP-YMQYCTKLWCTGKAKQWVCQTRHPFPWADGHSCEGKLC 501
 Db 466 DANROCCFTFGEEKSKHCPDAASTCTTLWCTGSGLLVCCTKHPPWADGHSCEGKWCVS 525
 QY 502 GACVERHNLNK-----RVDGSAKWDYPGCSRTCGGQVOLARQCTNPTIANGSKYCEGV 558
 Db 526 GKCVNKYTDK-KHFTATPHVSGWPGWGDSCRTCGGQVQTMRECDNPVEKNGGKYCEG 584
 QY 559 RVKVRONLEPCPSASGKSFREBQCEAFNGYHSTNRLTLAVAWVEKYSGVSPROCKKL 618
 Db 585 RVRVRSNIEDCPDN-NGKTFREBQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKL 643
 QY 619 ICRANGYGVFLVAPKVDGTLSPDSTSVGVGCKIKACGDNLSKKTIDKCGVCGGD 678
 Db 644 TCEAKGLGYFVLQPKVVDGTPCSPDSTSVGVGQCVKAGCDRIIDSKKFDKCGVCGGN 703
 QY 679 NKSCCKVTGLTFPMHGYNEVWAIAGASSIDIRQRYKELIGDDNYLALKNQSGKYLIN 738
 Db 704 GSTCKMKSGIVTSRPGYHDIVTIPAGATNIEVKHNRQSRNNGSEFLAIRAADGYIILN 763
 QY 739 GHEVVSVERDLVVKSLRYSGTGAVESLQASRPILPTVVEVLSVGKMTPPRVRYSF 798
 Db 764 GNFTLSTLEQDLYTKGVLRVSGSSAALERISFSLKPELTITQVLMVGHALRPKIKFTY 823
 QY 799 YLPKEPREDKSSHPPKDPGRGPSVLHNSVLSLSNQVEQDPPARWVAGSWGPCASCGSG 858
 Db 824 FMKKKTES-----FNAIPTFS-----EWVIEENGESKTCGSG 856
 QY 859 LQKRVDCRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSKCGRGFOR 915
 Db 857 WQRRVVOCRDINGH--PASECAKEVKPASTRPCADLPCHMQVQWDSWPCSKTCGKYKK 913
 QY 916 R-SLKCVGHGGRLLARDQCNLHKKPQE-LDFCVLRFC 950
 Db 914 RCLKCVSHDGGVLSNESCDELKPKHYIDFCTLTQC 950
 RESULT 14
 AAB21265
 ID AAB21265 standard; Protein; 896 AA.
 XX
 AC AAB21265;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Mouse metalloproteinase ADAMTS-1.
 XX
 KW Mouse; ADAMTS-1; metalloproteinase; ADAM;
 KW a disintegrin and metalloproteinase domain; thrombospondin domain;
 KW vaccine; neutropic; neuroprotective; antiparkinsonian;

KW cerebroprotective; cytostatic; antiarthritic; immunosuppressive;
 KW Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;
 KW autoimmune disease; brain tumour; brain injury.
 OS Mus musculus.
 PN W0200053774-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US06237.
 XX
 PR 08-MAR-1999; 99US-0264585.
 XX
 PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
 XX
 PI Kelner GS, Clark M, Maki RA;
 XX
 DR WPI; 2000-594326/56.
 XX
 PT Polynucleotide encoding novel members of a disintegrin,
 PT metalloproteinase and thrombospondin domain protein family used to
 PT prevent and treat Alzheimer's disease, cancer and autoimmune diseases -
 XX
 PS Disclosure; Fig 17; 129pp; English.
 XX
 CC The present sequence is mouse metalloproteinase ADAMTS-1. The
 CC ADAMTS family of proteins is closely related to the ADAM (A Disintegrin
 CC and Metalloproteinase Domain) family. Members of the ADAMTS family
 CC contain a thrombospondin domain in addition to the disintegrin and
 CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are
 CC useful for the manufacture of medicaments for treating conditions
 CC associated with neuroinflammation and/or neurodegeneration, such as
 CC Alzheimer's disease, Parkinson's disease and stroke. They are also
 CC useful for treating conditions associated with cell proliferation, cell
 CC migration, inflammation and/or angiogenesis, such as cancer, arthritis
 CC and autoimmune diseases. They can be used to treat patients afflicted
 CC with an invasive tumour, a brain tumour or brain injury.
 XX
 SQ Sequence 896 AA;

Query Match 45.0%; Score 2325; DB 21; Length 896;
 Best Local Similarity 48.1%; Pred. No. 2.4e-171;
 Matches 453; Conservative 144; Mismatches 235; Indels 110; Gaps 21;

QY 1 MLLLGILLAFAGRTAGG--SEPEREVVPIRLDPDINGRRYYWRGP-EDSGDGLIFQI 57
 Db 20 LLLLSITMLLCARGAHGRPTDEELVLP-SLE-----RAPHGDSITTLRL--RL 66
 QY 58 TAFQEDFVILHTPDQAFLAPAFSTHGLGVPLQGLTGGS-----SDLRCFYSGHNA 109
 Db 67 DAFGOQLHLKLPDQSGFLAPGFTLQTV-----GRSPGSEAOHLDPDGLAHCFYSGIVNG 121
 QY 110 EPDSPAAYSLCGLRGAFGYGAEEYVISPLPNAS----APAAQRNSQGA----HLLRRG 161
 Db 122 DFGSAALSLCEGVARGVAFYLOGEEFFIQAPGVATERLAPAVPEESSARQFHILRR- 180
 QY 162 VPGGSGDPTSCGIVASGWNPAIRLALDPYKPRRAGFESRR- 206
 Db 181 -----ARGSGKACGVMD-----DETLPSTDSRSESQNTNQPVPFDPDTPQDAGK 226
 QY 207 -----SGRAKRFVSPRYVETLVVADESVMKFGADLEHYLLTLATARLVHPHSILNPI 262
 Db 227 SGPGSIRKRFVSSPRYVETLVVADQSMADFGSLGKHYLLTLFVSAARFKHFSIRNSI 286
 QY 263 NIVVVKVLLLRDRSGPKVTGNAALTNRNFCAMQKLNKVDKHEPYMDIAILFTRQDLC 322
 Db 287 SLVVVKILVIYEEQKGPVTSNAALTNRNFCWQKNSPSDRDPEDHYDTAILFTRQDLC 346
 QY 323 GATTCDTLGMADVGTMDPKSCSVIEDDGLPSAFTTAHELGHVFNPHNVKVCVEVFG 382
 Db 347 GSHTCDTLGMADVGTCDPSRSCSVIEDDGLOAQAAFTTAHELGHVFNPHDPAKHCASLNG 406

```
QY 383 KLRANHMSTLTIIDRANPWSACSAALITDFDSHGDPILLDOPSKPISELPEDLPGASY 442
DB 407 VTGDSHLMASMLSSLDSPWSPCSAYMTSFLDNHGHGECIMDKFQNPILPSPDLPGTLY 466
QY 443 TISQCELAAGVSGPCPCP-YMOYCTKLWCTGKAGQWQTRHFWADGTSCEGKGLCK 501
DB 467 DANRQCOFTFGEESKPCPDAASTCTTLWCTGTSGLLVCTGTHFWADGTSCEGKWCVS 526
QY 502 GACVERHNLNKH---RVDSGNKAWDPYGPCSRTCGGGVQVARRQCTNPTFANGKYGCEV 558
DB 527 GKCVNKTDM-KHPATPVHSGWPGWGDSCRTCGGGVQVYTMRECDNPVFKNGKYGCEK 585
QY 559 RVKYSCLNLEPCSSASGSKSFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCKL 618
DB 586 RVIRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVETPKYAGVSPDRCKL 644
QY 619 ICRANGTYFYVLAPKVVDGTGLCSPOSTSVCVQKCIKAGCDNGLGSKKRFKCGVCGGD 678
DB 645 TCEAKGIGYFVLQPKVVDGTGCPSPDSTSVCVQKQCVKAGCDRIIDSKKKFCKGVCNGN 704
QY 679 NKSCKKVTGLFTKPMHGYNEFVVAIPAAGSIDIRQYKGLIGDDNYLALKNSQGYLLN 738
DB 705 GSTCKMGSIVTSTRPGYHDIVTIPAGATNIEVKHNRQGSNRNGSFLAIRADGTYYILN 764
QY 739 GHFVSAVERDLVVGSLRYSGTGTAVESLQASRPILPTEVLVSGKMTTPPRVRYSF 798
DB 765 GNFTLSLEQDLYKGTVLRYSGSSAALERIRSFSPLEPLTIQVLMVGHALRPKIKFTY 824
QY 799 YLPKPEREDKSHPKDRGSPSVLHNSVLSNOVEQDPPARPWAGSGKPCASCGSG 858
DB 825 FMKKKTES-----FNAIPTFS-----EWVIEKGECSKTCGSG 857
QY 859 LOKRAVDCRSAGORTVPACDAAH--RPVETOACGE-PCPTW 897
DB 858 WQRRVQCRDINGH---PASECAKEVKPASTRPCADLPCEPHW 896

RESULT 15
AAW78435
ID AAW78435 standard; Protein; 727 AA.
AC AAW78435;
XX
XX
DT 11-MAY-1999 (first entry)
XX Human ADAMTS-1 protein.
XX Metalloproteinase-disintegrin protein with a thrombospondin domain; ARDS;
KW ADAMTS-1; drug composition; foodstuff; leukocyte; thrombocyte; hepatitis;
KW blood; erythrocyte; inflammatory disease; rheumatoid arthritis; asthma;
KW nephritis; Crohn's disease; acute respiratory disease syndrome.
XX Homo sapiens.
XX
XX W09855643-A1.
XX
XX 10-DEC-1998.
XX
XX 03-JUN-1998; 98WO-JP02449.
XX
XX 03-JUN-1997; 97JP-0160422.
XX
XX (KURE ) KUREHA CHEM IND CO LTD.
XX
XX Hakozaaki M, Hirose K, Inoguchi E, Ishida Y, Ishioka K;
XX Kuno K, Matsushima K;
XX WPI; 1999-070277/06.
XX
XX N-PSDB; AAX17990.
XX
XX Human metalloproteinase-disintegrin protein with thrombospondin
XX domain - useful as leukocyte and thrombocyte decreasing and
XX erythrocyte increasing agent
```

```
XX
PS
XX
XX This sequence represents a novel human metalloproteinase-disintegrin
XX protein with a thrombospondin domain (ADAMTS-1). The protein may be used
XX in drug compositions and foodstuffs, as an agent for decreasing the
XX leukocyte and thrombocyte blood count and increasing the erythrocyte
XX blood count, e.g. for treatment of inflammatory diseases such as
XX rheumatoid arthritis, hepatitis, nephritis, Crohn's disease, asthma
XX and ARDS.
XX
XX Sequence 727 AA;
XX
XX Query Match 44.1%; Score 2274; DB 20; Length 727;
XX Best Local Similarity 54.4%; Pred. No. 1.6e-167;
XX Matches 408; Conservative 129; Mismatches 173; Indels 40; Gaps 11;

QY 209 RAKRFVSIPRVEVTLVADSWKFGHADLEHLYLLTLATAARLYRHPSTLNPINIVVVK 268
DB 9 RKRKFVSIPRYVETMLVADQSMAEFHGSLGKHLTLTLFSAARLYKRPSTNSVSLVVVK 68
QY 269 VLLLRDRSGPKVTGNAALTLRNFCAWQKLNKYSDKHPEYWDTAIFLTRQDLGATICD 328
DB 69 ILVHDEQKGEVTSNAALTLRNFCAWQKHNPPSDRDAEHYDTAILFTRQDLGSGQICD 128
QY 329 TLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHGLHGVFNPHDNPVCEVFGKLRANH 388
DB 129 TLGMADVTCMDPKRSCSVIEDDGLQAAFTTAHGLHGVFNPHDNPVCEVFGKLRANH 188
QY 389 WMSPTLIQIDRANPWSACSAALITDFDSHGDCILDOPSKPISELPEDLPASYYTLSSQC 448
DB 189 WNASMLNLDHSPWSPCSAYMTSFLDNHGHGECIMDKFQNPILPGLDPLGTLYDANRQC 248
QY 449 ELAFGVGSKPCP-YMOYCTKLWCTGKAGQWQTRHFWADGTSCEGKGLCKLRANG 507
DB 249 QFTFGEDSKHCPDAASTCTLWCTGTSGLLVCTGTHFWADGTSCEGKGLCKLRANG 308
QY 508 HNLNKH---RVDSGNKAWDPYGPCSRTCGGGVQVARRQCTNPTFANGKYGCEVRYKRS 564
DB 309 TD-RKHFDTPHGSGWPGWGDSCRTCGGGVQVYTMRECDNPVFKNGKYGCEVRYKRS 367
QY 565 CNLEPCSSASGSKSFREEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCKLICRANG 624
DB 368 CNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPDKCKLTQAKG 426
QY 625 TGYFVVLAPKVVDGTGLCSPOSTSVCVQKCIKAGCDNGLGSKKRFKCGVCGGDKNSCK 684
DB 427 IGYFFVLPQKVVDGTGCPSPDSTSVCVQKQCVKAGCDRIIDSKKKFCKGVCNGSGTCK 486
QY 685 VTGLTKPMHGYNEFVVAIPAAGSIDIRQYKGLIGDDNYLALKNSQGYLLNHFVVS 744
DB 487 ISGVS TSAKPGYHDIVTIPGATNIEVKHNRQGSNRNGSFLAIRADGTYYILNDYLS 546
QY 745 AVERDLVVGSLRYSGTGTAVESLQASRPILPTEVLVSGKMTTPPRVRYSFYLPKEP 804
DB 547 TLEQDIMYKGVLYRYSGSSAALERIRSFSPLEPLTIQVLMVGHALRPKIKFTYEVKKA- 605
QY 805 REDKSHPKDRGSPSVLHNSVLSNOVEQDPPARPWAGSGKPCASCGSGLOKRAV 864
DB 606 ---KES-----FNAIPTFS-----AWVIEWGECSKSLGQWRRLV 639
QY 865 DCRGSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSGRGRFORSLKCV 921
DB 640 ECRDINGQ---PASECAKEVKPASTRPCADHPQWOLGEWSSCKSTCGKYKKSLSACL 696
QY 922 GHGRLLRADQCNLHRRKPOE-LDFCVLRPC 950
DB 697 SHDGGVLSHESCDPLKPKHFDICTLTQC 726

Search completed: May 9, 2003, 15:22:18
Job time : 49 secs
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:18:36 ; Search time 15 Seconds
(without alignments)
2636.834 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLGLITLAFAGRTAGGSE.....DDCNLHRRKPOELDFCVLRKPC 950

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2486.5	48.2	967	1 ATSL_HUMAN	Q9uh8 homo sapien
2	2485.5	48.1	968	1 ATSL_MOUSE	P97857 mus musculus
3	2474	47.9	967	1 ATSL_RAT	Q9uq79 rattus norv
4	2138.5	41.4	890	1 ATSL_HUMAN	Q9uq79 homo sapien
5	2115	41.0	905	1 ATSL_MOUSE	P57110 mus musculus
6	1965	38.1	1629	1 ATSL_HUMAN	Q9p2n4 homo sapien
7	1915	37.1	837	1 ATSL_HUMAN	O75173 homo sapien
8	1856	36.0	930	1 ATSL_MOUSE	Q9r001 mus musculus
9	1847.5	35.8	930	1 ATSL_HUMAN	Q9una0 homo sapien
10	1745.5	33.8	630	1 ATSL_RAT	Q9esp7 rattus norv
11	1296	25.1	1077	1 AT10_HUMAN	Q9h324 homo sapien
12	1275	24.7	1593	1 AT12_HUMAN	P58397 homo sapien
13	1231	23.8	1211	1 ATSL_HUMAN	O95450 h adams-2
14	1215.5	23.5	997	1 ATSL_HUMAN	Q9ukp4 homo sapien
15	1192	23.1	1205	1 ATSL_BOVIN	P79331 b adams-2
16	1184.5	22.9	1205	1 ATSL_HUMAN	O15072 homo sapien
17	1048	20.3	860	1 ATSL_HUMAN	Q9ukp5 homo sapien
18	601.5	11.7	245	1 ATSL_BOVIN	Q9tt93 bos taurus
19	599	11.6	207	1 ATSL_BOVIN	Q9tt92 bos taurus
20	381.5	7.4	450	1 AT10_MOUSE	P58459 mus musculus
21	378.5	7.3	824	1 AD08_HUMAN	P78325 homo sapien
22	368	7.1	956	1 AD19_HUMAN	Q9h013 homo sapien
23	362	7.0	776	1 AD28_MACFA	O9xsl5 macaca fasc
24	349.5	6.8	813	1 AD33_HUMAN	Q9bz11 homo sapien
25	343.5	6.7	571	1 DIS2_BOTJA	P30431 bothrops ja
26	342.5	6.6	774	1 AD28_MOUSE	Q9ukl6 mus musculus
27	340	6.6	775	1 AD28_HUMAN	Q9ukq2 homo sapien
28	337	6.5	857	1 AD22_MOUSE	Q9rlv6 mus musculus
29	334	6.5	920	1 AD19_MOUSE	Q9p0k1 mus musculus
30	332	6.4	906	1 AD22_HUMAN	Q9p0k1 homo sapien
31	324	6.3	903	1 AD12_MOUSE	Q61824 mus musculus
32	322	6.2	826	1 AD08_MOUSE	Q05910 mus musculus
33	314	6.1	909	1 AD12_HUMAN	O43184 homo sapien

RESULT 1

ID	ATSL_HUMAN	STANDARD;	PRT;	967 AA.
AC	Q9UH8; Q9UP80; Q9UH83; Q9P2K0; Q9NSJ8;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).			
DE	with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).			
GN	ADAMTS1 OR METH1 OR KIAA1346;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Casas C., Pritchard M.A., Estivill X., Arbones M.L.;			
RT	"Cloning, characterization and mapping on human chromosome 21 of the orthologue of murine Adamts-1.";			
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A., AND FUNCTION.			
RC	TISSUE=Heart;			
RX	MEDLINE=99367466; PubMed=10438512;			
RA	Vazquez F., Hastings G., Ortega M.A., Lane T.F., Oikemus S., Lombardo M., Iruela-Arispe M.L.;			
RT	"METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity.";			
RL	J. Biol. Chem. 274:23349-23357(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Endothelial cells;			
RX	MEDLINE=20247184; PubMed=10785405;			
RA	Glienke J., Schmitt A.O., Piliarsky C., Hinzmann B., Weiss B., Rosenthal A., Thierauch K.H.;			
RT	"Differential gene expression by endothelial cells in distinct angiogenic states.";			
RL	Eur. J. Biochem. 267:2820-2830(2000).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=20181126; PubMed=10718198;			
RA	Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes.XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro.";			
RL	DNA Res. 7:65-73(2000).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20289799; PubMed=10830953;			
RA	Hattori M., Fujiyama A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Park H.-S., Toyoda A., Ishii K., Taudien S., Blechschmidt K., Polley A., Ohki M., Takagi T., Sakaki Y., Lehmann R., Patterson D., Menzel U., Delabar J., Kumpf K., Lehmann R., Zimmermann W., Reichwald A., Rump A., Schilhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,			

ALIGNMENTS

34	307	5.9	789	1	AD07_RAT	Q63180 rattus norv
35	300	5.8	413	1	ACLA_AKAC	Q9pw35 agkistrodon
36	298.5	5.8	480	1	DISA_TRIGA	P15503 trimeresu
37	289	5.6	769	1	AD11_HUMAN	O75078 homo sapien
38	288	5.6	788	1	AD07_MOUSE	O35227 mus musculus
39	286	5.5	414	1	HRYD_CROAT	P15167 crotalus at
40	280	5.4	773	1	AD11_MOUSE	Q9rlv4 mus musculus
41	277.5	5.4	814	1	AD15_HUMAN	O13444 homo sapien
42	277.5	5.4	935	1	AD22_XENLA	O42596 xenopus lae
43	277	5.4	478	1	HRTE_CROAT	P34182 crotalus at
44	268	5.2	797	1	AD33_MOUSE	Q923w9 mus musculus
45	264	5.1	729	1	AD21_MOUSE	Q9ji76 mus musculus

RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharte M., Becken O., Desario A., Reichelt J., Kauter G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzaym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspou M.-L.,
RA "The DNA sequence of human chromosome 21.";
RA Nature 405:311-319(2000).
RL [6]
RL NATURE 405:311-319(2000).
RP SEQUENCE OF 418-967 FROM N.A.
RC TISSUE-Melanoma;
RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER (BY SIMILARITY). HAS ANGIOGENIC INHIBITOR
CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE.
CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1938-GLU-1-LEU-1939
CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; AF170084; AAF15317.1;
DR EMBL; AF060152; AAD48080.1; ALT_INIT.
DR EMBL; AF060152; AAF23772.1;
DR EMBL; AB037767; BAA92584.1; ALT_INIT.
DR EMBL; AP001697; BAA9502.1;
DR EMBL; AL162080; CAB82413.1;
DR MEROPS: M12.222;
DR Genew; HGNC:217; ADAMTSL.
DR MIM; 605174;
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000130; Zn_MTpeptidse.
DR Pfam; PR00090; tsp_1; 6.
DR Pfam; PF01421; Reprolysin; 2.
DR Pfam; PF01562; Pep_M12B_propep; 2.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS00427; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS00427; ZINC_PROTEASE; 1.
DR PROSITE; PS00427; ZINC_PROTEASE; 1.
KW Hydrolyase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL
FT 1 49
FT POTENTIAL.
FT FT
FT PROPEP 50 252
FT CHAIN 253 967
FT ADAMTS-1.
FT SITE 198 198
FT ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 401 401
FT BY SIMILARITY.
FT ACT_SITE 402 402
FT ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 405 405
FT ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 411 411
FT DISINTEGRIN-LIKE.
FT DOMAIN 476 559
FT TSP TYPE-1 1.
FT DOMAIN 560 616

FT	DOMAIN	617	724	CYS-RICH.
FT	DOMAIN	725	849	SPACER.
FT	DOMAIN	850	908	TSP TYPE-1 2.
FT	DOMAIN	909	967	TSP TYPE-1 3.
FT	DOMAIN	943	946	POLY-LYS.
FT	DOMAIN	947	947	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	547	547	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	720	720	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	764	764	P -> A (IN REF. 4 AND 5).
FT	CONFLICT	227	227	O -> H (IN REF. 1).
FT	CONFLICT	468	468	S -> N (IN REF. 1).
FT	CONFLICT	561	561	
FT	SEQUENCE	967	AA: 105383 MW; C189389324741ED1 CRC64;	
QY	Query Match	48.2%;	Score 2486.5; DB 1; Length 967;	
QY	Best Local Similarity	49.0%;	Pred. NO. 1.2e-167;	
QY	Matches	486;	Conservative 154; Mismatches 250; Indels 101; Gaps 24;	
QY	1	MLLLGLTLTAFAGRTAGSGEPERVVRIRLDPDINGRRYWRKPEDSGDQGLFIQITAF	60	
QY	36	LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTRRLRLHAF	81	
QY	61	QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQSLTGSSDLRRCFYSGDVNAEPDSF	114	
QY	82	DOQLDLRLPDSFLAPGFTLQNGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA	136	
QY	115	AAVSLCGGLRGAFYGAEGYVVISPLPNAS---APAAQRNSQGA---HLLQ---ERGVPGG	165	
QY	137	AALSCLGVRGAFYLLGEAYFTQPLPAASERLATAAPGKPPAPLOFHLRLNRNRODVG	196	
QY	166	PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRRAGFSGESRRRS	207	
QY	197	TCGVWDEPRPTKAEDEDETEGEDEGPQWS-----PQDPALOGVGP-TGTGS	247	
QY	208	GRAKRFVSPRYVETLVVADESMVKFADLEHLYTLTAARLYRHPISILNINIVV	267	
QY	248	IRKKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTFSVAARLYKHPISIRNSLVVV	307	
QY	268	KVLLLRDRDSGPVGTGNAALTENFCAWOKLANKVSDKHPEYWDVTAIFTRDOLCGATTC	327	
QY	308	KILVIHDEQKGEVTSNAALTENFCAWOKLANKVSDKHPEYWDVTAIFTRDOLCGATTC	367	
QY	328	DTLGMADVGTWCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPMHDMVKCEVEFGKLRAN	387	
QY	368	DTLGMADVGTWCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPMHDMVKCEVEFGKLRAN	427	
QY	388	HMSPTLIQIDRANPWSACSAIITDPLDSGHGCLLDQPSKPLSLPELPGASYTLSCQ	447	
QY	428	HMASMLSLNLDHSPQSPSCSAYMITSFLDNGHGECLMDKPNPIQLPGDLPGETYDANRQ	487	
QY	448	CELAFGVGSKPCP-YMOYCTKLCWCTGKAKGOMVCOYTRHFPWADGTSCGEGKCLKGACVE	506	
QY	488	CQFTFGEDSKHCPDAASTCTLWCTGTSGGVLYCQTKHFPWADGTSCGEGKCLKGACVE	547	
QY	507	RHNLNKH---RVDGSAWAKWDPYPCSRCTCGGGVOLARRCQTNPTTPANGKYCGVRYR	563	
QY	548	KTD-RKHEDTTFHSGWCMGMPGDCSTCGGGVQYTRMRECDNFPVKNKGKCYGKRYR	606	
QY	564	SCNLEPCPSSASGKSFREOCEAFNGYHNHNTLTLAVANVPKYSGVSPDKCKLICRAN	623	
QY	607	SCNLEPCPSSASGKSFREOCEAFNGYHNHNTLTLAVANVPKYSGVSPDKCKLICRAN	665	
QY	624	GTGVFYVLAPKVDGTLCSPDSTSVCGVQGCIRKAGCDNGLSKKRFKDCGVCVGGDNKSK	683	
QY	666	GIGYFVVLAPKVDGTLCSPDSTSVCGVQGCIRKAGCDNGLSKKRFKDCGVCVGGDNKSK	725	
QY	684	KVTGLTKPMHGYNFVVAIPAGASSIDIRQGYKGLICDDNYLALKNSQCKYLLNGHFV	743	
QY	726	KISGVSITSAKPYHDIITPTGATNIEVKQRNORSGNRNNGSFLAKAADGTYLLNGDYTL	785	
QY	744	SAVERDLVWKSLLRYSGTGAVERSLQASRIPLVPLVSVLGVKMTPPRVKYSFYLPKE	803	
QY	786	STLEDIMYKGVVLRYSGLSSAALIRISFSPLKEPLTIQVLTGNALRPKIKYFYEVKKK	845	

QY 804 PREDKSSHPKDPGSPVLSVLSNQVEQPDPRPARVAGSWGSPCSASGSGLQKRA 863
 DB 846 ---KES-----FNAITFS-----ANVIEENGECSEKSCWQRRRL 878
 QY 864 VDCRGSAQRTVPACDAH--RPVETQAGE-PCPTWELSAWSPCKSCGRGFQRRSLKC 920
 DB 879 VECRDINGQ---PASECAKVKFPASTRPCADHPCPQQLGEWSSCKTCGKGYKKRSLKC 935
 QY 921 VGHGRLLRDQCQNLHKKPQE-LDFCVLRPC 950
 DB 936 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966

RESULT 2

ATSI_MOUSE STANDARD; PRT; 968 AA.

AC P97857; 054768;
 30-MAY-2000 (Rel. 39, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
 GN ADAMTS1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RA MEDLINE=98110583; PubMed=9441751;
 RX Kuno K., Lizasa H., Ohno S., Matsushima K.;
 RT "The exon/intron organization and chromosomal mapping of the mouse
 ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";
 RL Genomics 46:466-471(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97150761; PubMed=8995297;
 RX Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,
 RA Matsushima K.;
 RT "Molecular cloning of a gene encoding a new type of metalloproteinase-
 disintegrin family protein with thrombospondin motifs as an
 inflammation associated gene.";
 RL J. Biol. Chem. 272:556-562(1997).
 [3]
 RP CHARACTERIZATION, AND MUTAGENESIS OF GLD-403.
 RX MEDLINE=99303657; PubMed=10373500;
 RA Kuno K., Terashima Y., Matsushima K.;
 RT "ADAMTS-1 is an active metalloproteinase associated with the
 extracellular matrix.";
 RL J. Biol. Chem. 274:18821-18826(1999).
 [4]
 RP FUNCTION.
 RX MEDLINE=20389568; PubMed=10930576;
 RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,
 RA Ohno H., Matsushima K.;
 RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";
 RL FEBS Lett. 478:241-245(2000).
 [5]
 RP FUNCTION, AND INDUCTION.
 RX MEDLINE=20243757; PubMed=10781075;
 RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
 RA Richards J.S.;
 RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and
 cathepsin L proteases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
 CC INVOLVED IN ITS TURNOVER, HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
 CC SIMILARITY). ACTIVE METALLOPROTEINASE, WHICH MAY BE ASSOCIATED WITH
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
 CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
 CC SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1691-GLU-|-LEU-1692

CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
 CC COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
 CC MATRIX.
 CC -!- INDUCTION: INDUCED IN VITRO IN COLON ADENOCARCINOMA CELLS BY
 CC INTERLEUKIN-1, OR IN VITRO IN KIDNEY AND HEART BY
 CC LIPOPOLYSACCHARIDE. ALSO INDUCED BY LH STIMULATION IN GRANULOSA
 CC CELLS OF PREOVULATORY FOLLICLES.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A PURIN ENDOPEPTIDASE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
 CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
 CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 7.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB001735; BAA24501.1; ALT INIT.
 CC EMBL: D67076; BAA11088.1; ALT_FRAME.
 CC MEROPS: M12.222; -.
 CC MGD: MGI:109249; Adamts1.
 CC InterPro: IPR001762; Disintegrin.
 CC InterPro: IPR002870; Pep_M12B_propep.
 CC InterPro: IPR001590; Reprolysin.
 CC InterPro: IPR000884; TSPI.
 CC InterPro: IPR000130; Zn_MTPeptase.
 CC Pfam: PF00090; tsp_1; 3.
 CC Pfam: PF01421; Reprolysin; 1.
 CC Pfam: PF01562; Pep_M12B_propep; 1.
 CC SMART: SM00209; TSPI; 3.
 CC PROSITE: PS00215; ADAM_MEPRO; 1.
 CC PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 CC PROSITE: PS00092; TSPI; 3.
 CC PROSITE: PS00142; ZINC_PROTEASE; 1.
 CC Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;
 CC Repeat; Extracellular matrix; Heparin-binding.
 CC SIGNAL 1 48 POTENTIAL.
 CC PROPEP 49 253
 CC CHAIN 254 968
 CC SITE 206 206
 CC METAL 402 402
 CC ACT_SITE 403 403
 CC METAL 406 406
 CC METAL 412 412
 CC DOMAIN 477 559
 CC DOMAIN 560 617
 CC DOMAIN 618 725
 CC DOMAIN 726 850
 CC DOMAIN 851 909
 CC DOMAIN 910 968
 CC DOMAIN 195 199
 CC CARBOHYD 548 548
 CC CARBOHYD 721 721
 CC CARBOHYD 765 765
 CC CARBOHYD 783 783
 CC CARBOHYD 946 946
 CC MUTAGEN 403 403
 CC CONFLICT 335 335
 CC CONFLICT 425 425
 CC SEQUENCE 968 AA; 105841 MW; 42EBDA55499FB6C1 CRC64;

Query Match 48.1%; Score 2485.5; DB 1; Length 968;
 Best Local Similarity 48.0%; Pred. No. 1.4e-167;
 Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

```
QY 1 MLLIGILTAFAGTAGG--SEPEREVVVVIRLDPDINGPRYYWRGP-EDSGDQGLIFQI 57
Db 37 LLLLASITMLCARAGHPTEDELVP-SLE-----RAPHGDSITTLRL--RL 83
QY 58 TAFQEDFYHLPTDAQFOLAPAFSTHGLVPGVQGLTGG--SDLRKCFYSGDVNA 109
Db 84 DAFGQQLHLKLQPDGSLFAPGTLQTV-----GRSPGSEAQHLDTGLDHLAHCYSGTVNG 138
QY 110 EPDSPAANVSLCGLRGARGYGAEVISPLPNAS-----APAAQNSQGA-----HLLQBRG 161
Db 139 DFGSAALSLCEGVGAFYLOGEFFIOPAPGVATERLAPAVEESARPOFHILRRR- 197
QY 162 VFGPGSGDPTSCGVASGNPAILRALDPYKPRRAGFGESESRRR----- 206
Db 198 -----RGSGAGCGVMD-----DETLPDSRPESONTRNQWVPVROPTPDACKP 243
QY 207 -----SGRAKRFVSIPIRYVELTVVAVDESVMKFGADLEHYLLTLATARLYRHPSILNPI 262
Db 244 SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSLKHLYLLTLFSAARFYKHPISRNSI 303
QY 263 NIWVKVILLRDBSGPKVTGNAALTIRNFCWOKKLVSDKHPEYNDTALLFTRODLC 322
Db 304 SLVVKILYIEEQKGPEVTSNAALTIRNFCWQKHNSPDRDEHYDTALLFTRODLC 363
QY 323 GATTCTDLGADVGTMCDDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNVNVKVCVEYFG 382
Db 364 GSHTCTDLGADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKHCAASLNG 423
QY 383 KLRANHMGSPTLIQIDRANPWSACSAITDFDLSHGDCILDDQSPKPLSPEDLPQASY 442
Db 424 VTGDHSLMASLSSLDHSPQSPCSAYMVTSLFDNGHGECLMDKPNPKLPSDLPGTLY 483
QY 443 TLSQCELAGFVGSKPCP--YMOYCTKLWCTGKAGKGMVCTOHRFFWADGTSCEGKLCIK 501
Db 484 DANQCOFTGEESKCKCPDAASTCTLWCTGTSGLLVCQTKHFPWADGTSCEGKWCVS 543
QY 502 GACVERHNLNKH---RVDGSAKWDPYGPCSRCTCGGGVGLARQCTINPTTPANGKYCEGV 558
Db 544 GKCVNKTDM-KHPATPVHSGWPGWGDGCSRTCGGGVQYTMRECDNPFVKNKGKYCEGK 602
QY 559 RVYRSCNLEPCSSASGSKSFREOCEAENGYNHSTNRLTLAVANPKYSGVSPDKCKL 618
Db 603 RVYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKSGFNEPTVEWTPKYAGVSPDKCKL 661
QY 619 ICRANGTYPIVLAPKVDGTLCSPDSTSVYQGGKIKAGCDNLSGKKRFDKCGVCGGD 678
Db 662 TCEAKGIGYFVLQPKVDGTPCSPDSTSVYQGGVQCVKAGCDRIILSKKKFDKCGVCGGN 721
QY 679 NKCKKVTLFTKPMHGYNEFVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQGYLLN 738
Db 722 GSTCKKMSGIVTSTRPGYHDIVIPAGATNIEVKHNRQSRNNGSFLAIRAADSTYILN 781
QY 739 GHFVSAVERDLVVKSLRYSGTGTAVESLQASRPILPLETVELVSGKMTPPRVRYSF 798
Db 782 GNFTLSLEODLYTKGVLYRSGSRALEIRSFSLKEPLTIOVLVGHALRPK.KFTY 841
QY 799 YLPKPREDEKSSHKDPGRGPSVLHNSVLSLNOVEQPDPRPARVAGWSGPCSACSGG 858
Db 842 FMKKTES-----FNAIPFS-----EWVTEEWGECSTKCGSG 874
QY 859 LQRAVDCRSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCRGFOR 915
Db 875 WQWVQVQCRDINGH---PASECAKEVKPASTRPCADLPCPHWQVGDWSPCSKTCGKGYKK 931
QY 916 RSLKCVGHGRLLARDOCNLRHXPQE-LDFCVLRPC 950
Db 932 RTLKCVSHDGGVLSNESCDPLKPKKHWDFTCLTQC 967
RESULT 3
ATSL_RAT
ID_ATSL_RAT STANDARD; PRT; 967 AA.
AC Q9WUQ1; Q9ER11;
```

```
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN ADAMTS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
RA Little S.P.;
RT "Induction of a disintegrin and metalloproteinase with the
thrombospondin type I motif (ADAMTS).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 18-967 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=20304099; PubMed=10847486;
RA Diamantis I., Luehthi M., Hoesli M., Reichen J.;
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
endothelial cells in cirrhotic rats.";
RL Liver 20:165-172(2000).
CC -|- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
SIMILARITY).
CC -|- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1683-GLU-1-LEU-1684
SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
CC -|- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (by similarity).
CC -|- INDUCTION: DOWN-REGULATED IN ENDOTHELIAL CELLS DERIVED FROM
CIRRHOTIC LIVER.
CC -|- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -|- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
SIMILARITY).
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -|- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -|- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF149118; AAD34012.1; -.
DR EMBL; AF304446; AAG29823.1; -.
DR MEROPS; M12.222; -.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000130; Zn_M12peptidse.
DR Pfam; PF00090; tsp_1; 6.
DR Pfam; PF01421; Repolysin; 2.
DR Pfam; PF01562; Pep_M12B_propep; 2.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Extracellular matrix; Heparin-binding.
```

FT SIGNAL 1 54
 FT PROPEP 55 252
 FT CHAIN 253 967
 FT SITE 205 205
 FT METAL 401 401
 FT ACT_SITE 402 402
 FT METAL 405 405
 FT METAL 411 411
 FT DOMAIN 476 558
 FT DOMAIN 559 615
 FT DOMAIN 616 724
 FT DOMAIN 725 857
 FT DOMAIN 858 907
 FT DOMAIN 908 967
 FT DOMAIN 194 198
 FT CARBOHYD 547 547
 FT CARBOHYD 720 720
 FT CARBOHYD 764 764
 FT CARBOHYD 782 782
 FT CARBOHYD 945 945
 FT CONFLICT 21 21
 FT CONFLICT 26 31
 FT CONFLICT 49 49
 FT CONFLICT 72 72
 FT CONFLICT 79 79
 FT CONFLICT 249 249
 FT CONFLICT 262 265
 FT CONFLICT 607 607
 FT CONFLICT 936 936
 FT CONFLICT 962 962
 FT CONFLICT 967 967
 FT SEQUENCE 967 AA; 105705 MW; F93C864F6DCB4CF CRC64;

Query Match 47.9%; Score 2474; DB 1; Length 967;
 Best Local Similarity 48.4%; Pred. No. 9,2e+167;
 Matches 477; Conservative 162; Mismatches 255; Indels 92; Gaps 22;

QY 1 MLLGILLTAPAGTAGG--SEPEREVVPIRLDPDINGRRYYRGHEDSDQGLIFQIT 58
 DB 37 LLLLASITMLCVGAHGRTEDEELVL-----PSLERARGH-----DSTILLRLD 83
 QY 59 AFOEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGG--SRLRRCFYSSDVNAE 110
 DB 84 AFGQQLHLKLPQDSFLAPGTLQTV-----GRSPGEAHLDPGLAHCFYSGTVNGD 138
 QY 111 PDSFAVSLCGLRGAFGYGAEYVISPLPNAS-----APAAQRNSQA-----HLL--QRR 160
 DB 139 PSSAALSILCEGVRAFYLQGEFFIQAPAVATERLWPAEPKEESJAPPRFHILRRRR 198
 QY 161 GVPGFSGD-----PTSRCGVASGNPAIRLALDPKPRAGFGRSRRRSGRAKRFV 214
 DB 199 GSGGAKCGVMBDETLPTNSGRESQNTPDQWPLRNP-TPQAG--KPTGPGSIRKRFV 254
 QY 215 SIPRVETLVVADESVMKFGHADLEHLLTLLATARLYRHPSTILNINIVVKKVILLRD 274
 DB 255 SSPRVETLVVADQSMADFHGSLGKHLTLTFSVAARFYKHPSTNSLSLVVKKVILYIE 314
 QY 275 RDSGPKVTGNAALTRNFCARWKLNVSKHPEYWDTAIFLTRQDLICGATTCDTLGNAD 334
 DB 315 EQKGPEVTSNAALTRNFCQWQKQNSPSDRDPEHYDTAIFLTRQDLICGSHCTDLGNAD 374
 QY 335 VGTMCDPKRRSSVIEDDGLPSAFTAHGLHGVFNPNHNDVNVKVEEFGKLRANHMSPFL 394
 DB 375 VGTGCDPSRSCSVIEDDGLQAAFTAHGLHGVFNPNHNDHDAKAGFNGVSGDSHLMASML 434
 QY 395 IQIDRANPWSACSAIITDFLDLGHGCDLLDQPSKPTSLPDLFGA-YTLSQCELAAGV 454
 DB 435 SSLDSHQSPWSPCSAYMTVTFSLDNHGECMLDKPQNPILKPSDLPTGLYDANRQQQFTGE 494
 QY 455 GSKPCP-YMQVCTKLWCTGKAGQWVCOTRHFPPWADGTSCEGSKLCLKAGCVERHNLKH 513
 DB 495 ESTHCPDRASTCLWCTGTSGLLVCQTKHFPWADGTSCEGKWCVKCNKIDM-KH 553
 QY 514 ---RVDGSWAKWDPIGPCSRTCGGSQLARRQCTNPTPANGGKYCEGVRYKRYSCNLEPC 570

DB 554 FATPVHGSWGPWGPWGDSCSRTCGGVQYTMRECDNPVKNNGRYCEGKRYRYRSCNIEDC 613
 QY 571 PSSASGKSPREOCEAFNGYNHSTNRLTLAVAVPKYSGVSPRDKCLICRANGTCYFV 630
 DB 614 PDN-NGKTFREOCEAHNEFSAFNEPTVETPKYAGVSPKDRKLTCEAKGIGYFFV 672
 QY 631 LAPKVVDDGTLCSPDSTSVCVQGVKICRAGCDNGLSGKRFRDKCGVCGDNKSKKVLLET 690
 DB 673 LQPKVVDDGTPCSPDSTSVCVQGVKAGCDRIIDSKKFKDKCGVCGNGSTCKKISUTVT 732
 QY 691 KPMHGINFVVAIPAGASSIDIRORYKYGKGLIGDDNYLAKNSQKYLNLNHFVVSAVREL 750
 DB 733 STRPGYHDIVTIPAGATNIEVKHNPGRSNGSFLAIRAADGTYILNGNFTLSTLEODL 792
 QY 751 VVKGSLRLYSGTGTAVESLQASRPILPLTVLSVSKGKMTPPRVRYSFYLPK--EPREDK 808
 DB 793 TYKGTVLRYSGSSAALERINRSFPLKEPLTIQVLMVGHALRPKIKTYIFMKKTEH---- 848
 QY 809 SSHPKDPRGFSVLHNSVLSLNQVEQPDPRPARVAGSNWGPCSASCGLQKRAYDCRG 868
 DB 849 -----FNAIPTFS-----EWVIEEWGECSTKCGSQWQRRVVECRD 883
 QY 869 SAGQIVPACDAAH--RPVETQACGE-PCPTWELSAWSCKSCGRCGFORRSUKCVHGG 925
 DB 884 INGH---PASECAKEVKPASTRPCADLPCEPRWQGVDMSPCKTCGKGYKRTKLCLSHDG 940
 QY 926 RLLARQCNLHRKPOE-LDFCVLRPC 950
 DB 941 GVLNSCSDPLKPKPKHYIDFCILTQC 966

RESULT 4
 AT58 HUMAN
 ID AT58_HUMAN STANDARD; PRT; 890 AA.
 AC Q9UP79; Q9NZS0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2) (METH-8).
 DE DE
 GN ADAMTS8 OR METH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=99367466; PubMed=10438512;
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S., Lombardo M., Iruela-Arispe M.L.;
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity";
 RL J. Biol. Chem. 274:23349-23357(1999).
 RN [2]
 RP SEQUENCE OF 195-440 FROM N.A.
 RX MEDLINE=20079168; PubMed=10610729;
 RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;
 RT "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";
 RL Genomics 62:312-315(1999).
 CC -!- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES.
 CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ADULT AND FETAL LUNG, LOWER EXPRESSION IN BRAIN, PLACENTA, HEART, STOMACH AND FETAL BRAIN AND KIDNEY.
 CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY

QY	177	ASGWNPAILRALDPYKPR	--RAGFGESSRR	-----RRS	207
Db	154	--WGAPARPL--PRGPWEVETGEGQERGDHQDSEEESEEEAGASPPPLPGAT	209		
QY	208	GRAKRFVSIPIRYVETILVWADSMVKFEGADLEHYLLTLATARLYRHPSILNPLNVVV	267		
Db	210	SRTKRFVSEARFVETLLVADASMAAFYGADLQNHILTLMSVAARIYKHPISKNSILNVV	269		
QY	268	KVILLDRDRSGPKVTGNAALTLRNFCAMQKLNKYSDKHPXYDWTAILFTRODLGCAT	-T	326	
Db	270	KVLIVEDIKMGPEVSDNGGLTLRNFCNWRFRNQPSDRHPHYDTAILTRQNFQCGEGL	329		
QY	327	CDTLGADWAGVCTMCDPKRCSVTEDDGLPSAFATTAHELGHVFNHMDNVKVCVEEFGKLR	386		
Db	330	CDTLGADWAGITICDPNKSCSVIEDEGSQAHTLAHELGHVLSMPHDDSPCTRLTFPGMCK	389		
QY	387	NHMSPTLIQIDRANPWSACSAAITDFLDSHGDCLLDQPSKPISLPEDLPG--ASYTL	444		
Db	390	HHVMAPLFVHLNQLTPWSPCSAMYLTELLDGGHDCLLDAPGAALPLPTGLPCRMAIYQL	449		
QY	445	SQCELAFAVGVSKEPCP---YMQYCTKLWCTGKAKGMQWQCTRH--FPWADGTSCGEGKLC	499		
Db	450	DOQCROIFGDFRHCPTNTSAQDYCAOLWCHTDG-APPLCHTNGSLPWADGTPCGPGHLC	508		
QY	500	LKACAYVERHNLNKR--VDGSMAKWDPYPCSTRCTGGGVQLARRQCTNPTPANGKYCEG	557		
Db	509	SEGSCLPEEVEVERPKPVDDGNAPWGPWGSECTCGGVQFVSHRECKDPEPQNGGRYCLG	568		
QY	558	VRVYKSCNLEPCPSSGASKSPREOCEAFNGYNHSTNRLTLTAVAVPKYISGVSPRDCK	617		
Db	569	RRAKYOSCHTEEC--PDGKSFEQOCEKYNAYNYTMDGNL-LQWVPKYAGVSPRDCK	625		
QY	618	LICRANGTGYFVYLAIPKVDGTLCSPDSTSVCVOGCKIRAGCDGNLGSKKRDKCGVCGG	677		
Db	626	LFCRARGSEFKVFEAKVIDTGLCGPETLAI CVRGQCVKAGCDHVDSRPKLDKCGVCGG	685		
QY	678	DNKSCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQYKYGKGLIGDDNYLALKNSOGKYL	737		
Db	686	KGNSCRKVSGLTPTWYGVNDIYVTPAGATNIDVKQSHPGVQNDGNYLALKTADQGYLL	745		
QY	738	NGHFVVSAYERDLVVKGSLILRYSGTCTAVESLQASRPILPTVEVLSV-GKMTPPRYRY	796		
Db	746	NGNLAISAQDILVKGTLIKYSGSIATLRQSFRLPEPLTVQLTLVPGVEFPPKVKY	805		
QY	797	SFYLKPREDKSHPKDPRGSPVLNSVLSNQVEQPDPRPARWAGSGPCASGCG	856		
Db	806	TFVFPND--VDFSQMSKERRATTNIIQPLLH-----AQWLGDWSECSSTCG	850		
QY	857	SGLQKRAVDCRGSAGQRTVPACDAARHPVETQAC-CEPCP	895		
Db	851	AGWQRTVECRDPSGOASA-TCNKALKPEDAKPCESQLCP	889		
RESULT 5					
AT58_MOUSE					
ID	AT58_MOUSE	STANDARD;	PRT;	905 AA.	
AC	P57110;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase				
DE	with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).				
GN	ADAMTS8.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE 2007919166; PubMed=10610729;				
RA	Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;				
RT	"ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on				
RT	mouse chromosome 9 and human chromosome 11".				

RA Maki R.A.;
RT "ADAMS 9, a novel member of the ADAM-TS/Metallospandin gene
family";
RL Genomics 67:343-350(2000).
RN [2]
RP SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).
RX TISSUE=Brain;
RC MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
The complete sequences of 150 new cDNA clones from brain which code
for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.
CC EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,
PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN
COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR
THYMUS.
CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
SIMILARITY).
CC -1- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 11 TSP TYPE-1 DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL: AF261918; AAF89106.1; -;
DR EMBL: AB037733; BAA92550.1; -;
DR HSPB: P15167.1;ATL.
DR MEROPS: M12.021; -;
DR MIM: 605421; -;
DR Genew: HGNC:13202; ADAMTS9.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Repolysin.
DR InterPro: IPR000684; TSP1.
DR InterPro: IPR000130; Zn_MTpeptidse.
DR Pfam: PF00090; tsp_1; 11.
DR Pfam: PF01421; Repolysin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR SMART: SM00209; TSP1; 12.
DR PROSITE: PS00215; ADAM_MEPRO; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE: PS00092; TSP1; 9.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Alternative splicing.
FT SIGNAL 1 18
FT PROPEP 19 287
FT CHAIN 288 1629
FT DOMAIN 509 587
FT DOMAIN 589 642
FT DOMAIN 645 752
FT DOMAIN 753 880
FT DOMAIN 999 1053
FT DOMAIN 1056 1108
FT DOMAIN 1111 1156
FT DOMAIN 1184 1239
FT DOMAIN 1240 1295
FT TSP TYPE-1 2.
FT TSP TYPE-1 3.
FT TSP TYPE-1 4.
FT TSP TYPE-1 5.
FT TSP TYPE-1 6.

FT DOMAIN 1332 1383
FT DOMAIN 1386 1439
FT DOMAIN 1445 1498
FT DOMAIN 1501 1554
FT DOMAIN 1562 1612
FT DOMAIN 1612 1662
FT DOMAIN 1662 1712
FT SITE 223 273
FT METAL 434 484
FT ACT_SITE 435 485
FT METAL 438 488
FT METAL 444 494
FT CARBOHYD 112 162
FT CARBOHYD 135 185
FT CARBOHYD 271 321
FT CARBOHYD 749 799
FT CARBOHYD 840 890
FT CARBOHYD 1213 1263
FT CARBOHYD 1267 1317
FT VARSPLIC 1084 1134
FT VARSPLIC 1073 1123
FT CONFLICT 367 377
FT CONFLICT F -> L (IN REF. 1).
SQ SEQUENCE 1629 AA; 182649 MW; C1C4CEFF58B941F CRC64;
Query Match 38.1%; Score 1965; DB 1; Length 1629;
Best Local Similarity 39.9%; Pred. No. 1.4e-130;
Matches 411; Conservative 145; Mismatches 351; Indels 122; Gaps 21;
QY 22 ERVVVPIRLDP-----DINGRRVYV-----RGPEDSGDGLFOITAFQ 61
DB 47 EYEIVSPIRVNAGLGEPTNVHFKTRRSINATDPWAFASSSSSSTSSQAHYRLSAFG 106
QY 62 EDFYHLTPDAQFLAPAFSTEHLGVV-----LQGLTGGSSDLRRCFYSGDVNAEPDSRAV 117
DB 107 QQPLENLNANAGFIAPLFTVLLGTPTGVNQTIFYSEEAELKCFYKGYVNTNSEHTAVI 166
QY 118 SLCCGLRGAFYRGABYVLSPLPNASAPAAORNSQGAHLQRRVGGPGSGDFTSRGVA 177
DB 167 SLCSGMLGTFRSHDGYFIEPLQSMDEQDEEQNKPHIYRRSAPQREPSTGRACOTS 226
QY 178 SGWNPAILRALDPYKPRRAGFGE-----SRSRRSRG 208
DB 227 EHKN---RHSKKKTRARKWGERINLAGDVAALNSGLATEAFSAVGKNTDNTREK:HR 283
QY 209 RAKRFVSIPIRYVETLVVADESVMKFGADLEHYLLTATAARLYRHPHSILNFINIVVK 268
DB 284 RTRKFLSYPRFEVLVVAADNRVSYHGENLQHYILTIMSIVASIYKOPSGIGNLILNIVVN 343
QY 269 VLLLRDRDSGPKVTGNAALTNRNCAWKKLNKVSCKHPEYWDATLFTRODLCA--TTC 327
DB 344 LIVIHNEQDGPISIFNAQTTLKNFCQWQHSKNSPGGIH---HDTAVLLTRQDICRAHDKC 400
QY 328 DTLMADVGTMCDPKRSVIEDDGLPSAFTTAHELGHVFNPHDPMNVKCYEEVEFGKLRAN 387
DB 401 DTGLAELGTICDPYRSCSISSEDSGLSTAFTAHLEGHVFNPHDPMNDDNNKCKEE-GVKSPQ 459
QY 388 HMMSPILQIDRANPWSACSAIITDFLSGCHDCLLDOP-SKPTSLPDLPGASYTLQ 446
DB 460 HVMAPTLNFTNPMWMSKCSRYITEFLDTGEGCLNPESESRPYPLVOLPGILYNVKNK 519
QY 447 QCLEAFGVGSKPCPYMOYCTKLWC---TGKAGQOMVQCTRHPPWADTSCGEGKCLCKGA 503
DB 520 QCELIPEGSGQVCPYMQCRRLMCNWNVGVHKG---CRTOHTFPWADTCECEPKHCKYGF 576
QY 504 CVERHNLNHRVDGSAKWDPYGPCSRTCGGVQVLAARQCTNPTPANGKYCEGVAVKYR 563
DB 577 CVPK-EMDVPVTDGWSGWSFGTCSRTCGGGIKTAIRECNRPKNGKCYCVGRMRKF 635
QY 564 SCNLEPCPSASGKSPREOCEAFNYSNINSTRNLTLAVAVPKYSGVSPDKCKLICRAN 623
DB 636 SCNTEPCLKQK--RDRFDEQCAHFDGKHFNGLLNPVRWPKYSGILMKRCKLFCRVA 693
QY 624 GTGYFYVLAPEVVDGTLCSLSPDSTSVCVQKCIKAGCDGNLGSKKRDKCGVCGGDNKSK 683


```

Db 694 GNTAYQLRDRVIDGTGCGODTNDICVQGLCRQAGCDHVLNSKARRDKCVGCGDNSSCK 753
QY 684 KVTGLFKPMHGVNFVVAIPAGASSIDIRQGVKGLIGDNDYALNLSQCKYLLNGHFVY 743
Db 754 TVAGTFNTVHGVNFVVAIPAGATNIDVROHSFSGTDDNDYALSSKGEFFLNGFVY 813
QY 744 SAVERDLVVKSLRYSGTGTAVESLQASRPIELTVEVLSVGKMTTPRVRYSFVLPKE 803
Db 814 TMAKREIRGNNAVVEYSGSETAVERINSTDRIEQELLQVLSVGKLYNFDVYSFNPIE 873
QY 804 PREDK---SSHPKDRGP-----SVLHNSVLSSNOVEQDPP--- 840
Db 874 DKPOQFYWNH-----GPMQACSKPCQGERKRLVCTRESQDLTVS---DQRCDRLPQPG 925
QY 841 -----ARWAGSWGPCSCGSGLOKRAVDCR-----GSAGQETVPACDAAR 883
Db 926 HITEPCGTCDLRRHWASRSECSAQGLGVRTLDIYCAKYSRLDGKTEKVDGFCSSHPK 985
QY 884 PVTQACGECPT--WELSAWSPCKSGRGFORRSKLCVGHGRLLRDQCNLHRKPOE 941
Db 986 PSNRKCSGBCNTGGWRYSAWTECSKCDGGTORRRAICVNRNDVLDKCT-HQEKVT 1044
QY 942 LDFCVLRPC 950
Db 1045 IQRCSEFP 1053

RESULT 7
AT54_HUMAN
ID AT54_HUMAN STANDARD; PRT; 837 AA.
AC Q75173; Q9UN83;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-4 precursor (PC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)
DE (ADMP-1).
GN ADAMTS4 OR KIA0688.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suvama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 5:169-176(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99286303; PubMed=10356395;
RA Tortorella M.D., Burn T.C., Pratta M.A., Abbaszade I., Hollis J.M.,
RA Liu R.-Q., Rosenfeld S.A., Copeland R.A., Decicco C.P., Wynn R.,
RA Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,
RA Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K.,
RA Hillman M.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,
RA Trzaskos J.M., Arner E.C.;
RT "Purification and cloning of aggrecanase-1: a member of the ADAMTS
RT family of proteins."
RL Science 284:1664-1666(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Sawaji Y., Nagase H., Saklatvala J., Clark A.R.;
RT "ADAMTS-4 genomic locus."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=20400518; PubMed=10827174;

```

```

RA Tortorella M.D., Pratta M.A., Liu R.-Q., Abbaszade I., Ross H.,
RA Burn T.C., Arner E.C.;
RT "The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for
RT aggrecan substrate recognition and cleavage."
RL J. Biol. Chem. 275:25791-25797(2000).
CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. COULD ALSO BE A ONE
CC CRITICAL FACTOR IN THE EXACERBATION OF NEURODEGENERATION IN
CC ALZHEIMER'S DISEASE.
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
CC site.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG AND HEART. EXPRESSED
CC AT VERY LOW LEVEL IN PLACENTA AND SKELETAL MUSCLES.
CC -!- INDUCTION: BY INTERLEUKIN-1.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FUIN ENDOPEPTIDASE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO AS ADAMTS2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
EMBL: AB014588; BAA31663.1;
EMBL: AF148213; AAD41494.1;
EMBL: AY04847; AAL02262.1;
MEROPS: M12.221;
GENE: HGNC:220; ADAMTS4.
MIM: 603876;
InterPro: IPR001762; Disintegrin.
InterPro: IPR001590; Reprolysin.
InterPro: IPR000884; TSP1.
InterPro: IPR000130; Zn_MTPeptidse.
Pfam: PF00090; tsp.1;
SMART: SM00209; TSP1;
PROSITE: PS00142; ZINC_PROTEASE; 1.
PROSITE: PS0215; ADAM_MEPRO; 1.
PROSITE: PS00092; TSP1; 1.
PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Extracellular matrix;
FT SIGNAL 1 51 POTENTIAL.
FT PROPEP 52 212
FT CHAIN 213 837
FT SITE 194 194
FT METAL 361 361
FT ACT_SITE 362 362
FT METAL 365 365
FT METAL 371 371
FT DOMAIN 437 519
FT DOMAIN 520 576
FT DOMAIN 577 685
FT DOMAIN 686 837
FT DOMAIN 247 252
FT CARBOHYD 68 68
FT CONFLICT 77 77
FT CONFLICT 626 626
FT CONFLICT 682 682
SQ SEQUENCE 837 AA; 90224 MW; 5DF9C9AC137DF41F CRC64;

```

Query Match

37.1%; Score 1915; DB 1; Length 837;

```
Best Local Similarity 46.3%; Pred. No. 2.1e-127;
Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

QY 1 MLLGILTLAFAGRTAGSGSEPERVVPVIRLDPDINGRRYYWRGPEUSGGGLFIQTAF 60
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 37 LLLLLALLSPLSARLASLPREEIVFPEKLNGSVL-----PGSGAPARLLCRLOAF 88
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 61 QDFVLHLTPDPAFLAPAFSTHGLVPLQGLTGSSDLRRCFYSGDVNAEPDSPAAYSLC 120
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 89 GTLLELEQDSGVQVEGLTVQYLGQAPL-LGGAEP--GYILGTINGDPESVASLHWD 145
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 121 GGL-GRGAFYRGAEYVISPPLNASAPAAQRNSQGAHLLOREGVPGSGDPTSCGVASG 179
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 146 GGALLGLVQYRGAEHLQPLEGGTPNSA--GGPGAHILRRK------SPASGGQPMCNV--- 196
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 180 WNPAILRALDPKPRACFGSRRSRGRKRFVSPRYVETLVVADESVMVKEFGADLE 239
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 197 -----KAPLGSFSPRPR--RAKRFASLSRFVETLVVADKMAAFHGAGLK 239
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 240 HYLLTLLATAARLYRHPSTLNPINIVYKVLRLDRDSDGPKVTGNAALTLRNFCAWOKKL 299
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 240 RYLLVMAAAKAFKPSIRNPVSLVTLVILGSGEESGQVGPSSAQTLSFCAMQRL 299
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 300 NKVSKHPEYWDTAILFTFRQDLGATCTDTLGMADVGMTCDPKRSQSVIEDDGLPSAFTT 359
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 300 NTPEDSDPHFDTAILFTFRQDLGATCTDTLGMADVGMTCDPKRSQSVIEDDGLQSAFTA 359
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 360 AHELGHVFNMPNDVKNVCEEVFGKL-RANHMSPTLIOIDRANPWSACSAALITDFLDSG 418
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 360 AHELGHVFNMLHDNPKCISLNGPLSTSRHVNAPVMAHVDEPWPSPCSARFITDFLDNG 419
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 419 HGCDLDDQSKPISLPEDLPKASYTLTSCQELAFGVGSKPCTM-QYCTKLWCTGKARGQ 477
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 420 YGHCLLDKPEALHLVPTFPKDYDADQCOLTFGPDSSHCPCLPPPCAAWLCSGHLNGH 479
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 478 MVCQTRHFWPADGTCGEGEKLCKGACVERHNLNKHVD--GSAKAWDPYGPCSCTCGG 535
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 480 AMCQTKHSPWADGTPCGPAQACMGRCGLHMDLODFENIPQAG-WGPGWPWGDSCSRTCGG 539
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 536 VOLARQCTNPTPANGKYCEGVYKRYSCNLEPCPSASGKFPREECEAFNGYNHSTN 595
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 540 VQFSRDCRTPVRNGKCYCEGRFRFRSCNTEDCP-TGSALTFRFEOCAA---YNNHRTD 595
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 596 RLTL---AVAWPKYSGVSPDKCKLICRANGTYFYVVLAPKVVDGTLCSPDSTSVQVG 652
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 596 LFKSFPQMDWPRTYTGVAPOQCKLTQARALGYVYVLEPRVVDGTPCSPDSSVCVQG 655
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 653 KTKACDGNLGSKKRFPDCKGCGGDKNSCKKVTGLTKPMHGYNFVVAIPAGASSIDIR 712
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 656 RCIHAGCDRIIGSKKFKDKCMVCGGDSGSGSKSGSFRKFRYGNVVTIPAGATHILVR 715
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 713 QRGYKGLIGDDNYLALKNSOGKYLNGHFVVSVERDLVVKGLS-LRYSGTGTAVESLQA 771
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 716 QQGNPG--HRSYIALKLPDGSYALNGEYTLMPSTDVVLPAVSLRSGATASETLSG 773
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 772 SRPILPPLAVEVLSVGKMTPTPRVRYSYFLPKPREDKSSHPKD 814
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 774 HGPLAQPLTLQVLVAGNPQDTRLRYSEFFVPRPTPTSPRPTPD 816

RESULT 8
ATSS_MOUSE
ID ATSS_MOUSE STANDARD; PRT; 930 AA.
AC Q9R001;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS) (Aggrecanase-2)
DE (ADMP-2) (Implantin).
GN ADAMTS5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99395124; PubMed=10464288;
RA Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;
RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RT zinc metalloproteinases.";
RL J. Biol. Chem. 274:25555-25563(1999).
CC -!- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN
CC PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
CC site.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN THE PERI-
CC IMPLANTATION PERIOD IN EMBRYO AND TROPHOBLAST AND AT LOW OR
CC UNDETECTABLE LEVEL THEREAFTER.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
EMBL; AF140673; AD56356.1; --
MEROPS; M12.225; --
DR MGD; MGI:1346321; Adamts5.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
KW SIGNAL
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 261 POTENTIAL.
FT CHAIN 262 930 ADAMTS-5.
FT SITE 209 209 CYSTEINE SWITCH (POTENTIAL).
FT METAL 410 410 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 411 411 BY SIMILARITY.
FT METAL 414 414 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 420 420 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 485 566 DISINTEGRIN-LIKE.
FT DOMAIN 567 623 TSP TYPE-1 1.
FT DOMAIN 624 731 SPACER.
FT DOMAIN 732 874 CYS-RICH.
FT DOMAIN 875 930 TSP TYPE-1 2.
FT DOMAIN 41 46 POLY-ALA.
FT DOMAIN 257 261 POLY-ARG.
FT CARBOHYD 498 498 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 728 728 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC... ) (POTENTIAL).
```

```
FT CARBOHYD 807 807 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 930 AA; 101780 MW; 84DB84B26170DADC CRC64;

Query Match 36.0%; Score 1856; DB 1; Length 930;
Best Local Similarity 43.2%; Pred. No. 3.4e-123;
Matches 370; Conservative 115; Mismatches 265; Indels 106; Gaps 13;

QY 90 GLTGSSDLRRCFYGDYNAEPDPSFAAVSLGGLRGAFYGAEVLSPLNASAPAAQR 149
Db 121 GLSASSGHRGHCYRGTVGDSRSLAVFDLCGLDGFVAVKHARYTKLPLRGSWAEYER 180
QY 150 -----NSOGAHLQRRG-----VPGPGSPDPTSRGCVASGWNPAILRALDPY 191
Db 181 IYDGGSSRLHVYNREGFSFEALPPRASCETPASPSPG----- 217
QY 192 KPRRAGFGESESRSSRG-----RAKRFYSIPRYVTVLTVAD 227
Db 218 -POESPSVHSRRRSALAPOLLDSAPSPGNAGPOTWNRRRRSISRAKQVLELLVAD 276
QY 228 ESMVKFHGADLEHYLLTLATAARLYRHPSILNPIVVKVLLLRDRDSGPKVTGNAAL 287
Db 277 SSMARYGRGLQHYLLTLASTANRLYSHASTENHRLAVKVVVLTDKDTSLEYSKNAAT 336
QY 288 TLRNFCAMQKLNKVDKHPYWDATLFTRODLGGATTCTDGLMADVGMCDPKRSCSV 347
Db 337 TLKNFCKWQHONQGLDDEHYDAALFTREDLCGHSCDITGLMADVGTICSPSCAV 396
QY 348 IEDDGLPSAFTAHGLHVFNMHPDNVCEVEFGKLRANHMSPILQIDRANPWSACS 407
Db 397 IEDDGLHAAFTVAHEIGHLGLSHDDSKFEENFGTTEDKRLMSILTSIDAKPWSKCT 456
QY 408 AAITDFLDGSHGCLLDQSPKSTLSPEDLPASYYTILSQOCELAFGSKPCPYMOYCTK 467
Db 457 SATITEFLDDHGNCILDLPRKQTLGPEELPGQTYDATQCNTLFGPEYSVCPGMDVCAR 516
QY 468 LWCTGKAKGQWCVOTRHPNADGTSCEGKLCCLKGACVER-----HNLNHRVDGSKAKW 522
Db 517 LWCVVRRQGMVCLTKLPAVEGTPCGKRGVCLQKCKVDTKKYKYSTSSH---GNWGSW 573
QY 523 DPGGPCSRTGGGVQOLARRCOTNPTPANGGKYGEGVVKYRSCNLEPCPSSASGKGFREE 582
Db 574 GPMQCSRSRSGGVQFAYRHCNPNAPRNSRGYCTGRKAIYRSCSVTPCP---PNGKFRHE 631
QY 583 QCEAFNCSYNSTRLLTAVAWPKYSVSPDRCKLICRANGTGYFYVLAPKVVVDGLTCS 642
Db 632 QCEAKNGYQSDAKGVKTFVEMWPKYAGVLPADVCKLTCRAKGTGYVWFSPKVTDTGECR 691
QY 643 PDSTSVCVQKCKIKAGDGNLGSKKRFDKCGVCGDNKCKKYTGTLFTKPMHGYNFVVAI 702
Db 692 PYSNSVCRGCVRTGCDGLIGSLQVLDKCGVCGDNSSCTKIIGTFNKKSGYTDVRI 751
QY 703 PAGASSIDIRQYKGLIGDDNIALKNSQKYLNLGHFVVSVAVERDLVVKYGLLRYSGT 762
Db 752 PEGATHIKVRQFRAKQDTPREPAYLALKKTKTEYLINGKYMISTEITIIDINGVMYSGW 811
QY 763 GTAVESI-----QASRILEPLFVLSVCKWTPPRVRSFYFLPKPREDKS----SHPKD 814
Db 812 SHRDDFLUGMGYSATKEI-----LIVQLATDPTKALGVRYSFEPKPKTKQVNSVISHGNS 868
QY 815 PRGPSVLHNSVLNSLNOVEQDPRPARWVAGSWGPGSCSGGLKRAVDCRGSAGORT 874
Db 869 KVGPI---HSTQL-----QWYTGFWLACSRCTDGTGWHTRVYQCO-DGNRKL 909
QY 875 VPACDAAHRRVETQAC 890
Db 910 AKGCLLSQRPSAPFKQC 925

RESULT 9
ATSS_HUMAN
ID ATSS_HUMAN STANDARD; PRT; 930 AA.
AC Q9UNAO; Q9UKP2;
DT 16-OCT-2001 (Rel. 40, Created)
```

```
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2)
DE (ADMP-2) (ADAM-TS 11).
GN ADAMTS5 OR ADMP2 OR ADAMTS11.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=99367476; PubMed=10438522;
RA Abbaszade I., Liu R.-Q., Yang F., Rosenfeld S.A., Ross O.H.,
RA Link J.R., Ellis D.M., Tortorella M.D., Pratta M.A., Hollis J.M.,
RA Wynn R., Duke J.L., George H.J., Hillman M.C. Jr., Murphy K.,
RA Wiswall B.H., Copeland R.A., Decicco C.P., Bruckner R., Nagase H.,
RA Ito Y., Newton R.C., Magolda R.L., Trzaskos J.M., Hollis G.F.,
RA Arner E.C., Burn T.C.;
RT "Cloning and characterization of ADAMTS11, an aggrecanase from the
RT ADAMTS family.";
RL J. Biol. Chem. 274:23443-23450(1999).
RN [2]
RS SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudo J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Deyand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [3]
RS SEQUENCE OF 413-930 FROM N.A.
RX TISSUE=Fetal brain;
RX MEDLINE=99395124; PubMed=10464288;
RA Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;
RA "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RA zinc metalloproteinases.";
RT J. Biol. Chem. 274:25555-25563(1999).
CC -!- FUNCTION: CLEAVES AGGECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC DESTRUCTION OF AGGECAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN
CC PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
CC site.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVEL, PRIMARILY IN PLACENTA
CC BUT ALSO IN OTHER TISSUES, SUCH AS HEART AND BRAIN, AND ALSO
CC CERVIX, UTERUS, BLADDER, ESOPHAGUS, RIB CARTILAGE,
CC CHONDROBLASTOMA, FIBROUS TISSUE AND JOINT CAPSULE FROM AN
CC ARTHRITIC PATIENT.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```



```
DR pfam: PF00090; tsp_1; 5.
DR pfam: PF01421; Reprolysin; 1.
DR pfam: PF01562; Pep_X12B_propep; 1.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; ZINC_PROTEASE; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR Hydrolase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT NON_TER 1
FT PROPEP <1 207 BY SIMILARITY.
FT CHAIN 208 1077 ADAMTS-10.
FT METAL 366 367 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 367 367 BY SIMILARITY.
FT METAL 370 370 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 376 376 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 434 520 DISINTEGRIN-LIKE.
FT DOMAIN 578 679 CYS-RICH.
FT DOMAIN 680 802 SPACER.
FT DOMAIN 521 577 TSP TYPE-1 1.
FT DOMAIN 799 860 TSP TYPE-1 2.
FT DOMAIN 862 918 TSP TYPE-1 3.
FT DOMAIN 922 976 TSP TYPE-1 4.
FT DOMAIN 981 1031 TSP TYPE-1 5.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 769 769 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1077 AA; 118072 MW; 3914DE18DCBFF587 CRC64;

Query Match 25.1%; Score 1296; DB 1; Length 1077;
Best Local Similarity 35.4%; Pred No. 1.3e-83;
Matches 335; Conservative 112; Mismatches 376; Indels 123; Gaps 32;

QY 24 EVVVPRLDPDINGR-----RYWGRPGDSGGQIFQITAFQEDFYHLHPDAQF 74
DB 14 EIAFPTRVDRH-NGALLAFPPPPRRQRGTGATAESRLFYKVASPSHFLLNLTSSRL 71
QY 75 LAPAFSTEHLGVPLQGLTGSSDLRCFYSGDYNAE-POSFAVSLCGGLRAGFYRGAE 133
DB 72 LAGHVSVEYW--TREGLAWRQARPHCLYAGHLQGGASSHVAISTCGGLHGLIVADEEE 129
QY 134 YVISPPLNASAPAAQRNSQGAHLQRPGVPGSPGDPTRSGCV--ASGW--NPAILRALD 189
DB 130 YLIEPLHGGPKGSRSPESGPHVYKRSLSLRHPLD--TAGGYRDEKPKWKGREWLRCLK 187
QY 190 PYKPRRAGFGESSRRRSRAKRFVSIPTVETLVVADESMMVKFHG-ADLEHYLLLLAT 248
DB 188 P--PPARPLGNETERGQPG-LKRSVSRYRYETLVVADKMMVAYHGRDVEQYVLAIMI 244
QY 249 AARLYRHSILNPINIVVKKVLLLRDSGPKVTGNAALTFRNCAWOKKL-----NK 301
DB 245 VAKLFQSSSGTNNILVTLRLITLTDQPLETHIAGKSLDSCFKWQKSIWNHSGHNA 304
QY 302 VSDRHPYWDTAIFLTRQDLC--GATTCDTLGMADVGTMDPKRSCSVIEDDGLPSAFT 359
DB 305 IPENGVAHDTAVLITRYDICIYKNKPCGTGLGLAPVGGMCERERSCSVNEIDIGLQAF 364
QY 360 AHELGHVFNPHDNVKKVEEVFG-----KLRAHMMSTLIQIDRANP--WSACSAALITD 413
DB 365 AHEIGHTFGMNHHDGVNCGSGARGODPAKLMAAHITMTK-----NPFVWSSCNRYITS 417
QY 414 FLDSGHGDCLLDQPSK-PISLPEDLPKASYTLSSQCEFLAFGVSGKPCPYMYQCTKLWCTG 472
DB 418 FLDGSLGLCLNRRPPRODFYVTPAPQADYADACQRFQHGVSROCKYGEVCSLWCLS 477
QY 473 KAKGMQVCQTRFFHWADGTSQGEKL-----CLKGACVVERINLNKRVDSGNKAWDPGPC 528
DB 478 KSNR---CITNSIPAAEGLTCQHTIDKGWCYKRVCP-FCSRPEGVGDGAWGPWPGDC 533

529 SRTCGGGVQLARRCQNTNPTPANGGKYCVGVVYKRSCLNLEPCPSSASGKSFREUQCEAFN 588
534 SRTCGGGVSSSRHDSRPRTIGGKYCLGERRRHRSCNTDDCPGGS--QDFREVQCSEFD 591
589 G-----YNHSTNRLTLAVAWPKYSGVSPRDKCKLICRANGTGYFVVLAPVVDGTLIC 641
592 SIPRGRFYKWKTYR-----GGGVK---ACSLTSLAEGFNFTYTERAAAVVDGTPC 638
642 SPDSSTVCVQKCIKAGCDGNLGSKKRFRDGCVCGGDNKSKKVTGLFT--KPMHYYNFV 699
639 RPDIVDICVSCECKHVGCDRLGDLREDKRCVCGDGSACETIEGVSPSPAGYEDV 698
700 VAIPAGASSIDIIRORYKGLIGDDNYLALKNQSKYLLNGHVFVSAVERDLVVK ISLLRY 759
699 VWIPKGVSHFIQD-----LNLSLSHLALKGDOESLLEGLPTGPQPHR-LPLAGTTFQL 752
760 SGTGTAVESLQASRPILPTLTVLSVGMTPPRVRYSFYLPKPREDKSSHEKDPGPGS 819
753 RQGPQVQSLEALGPTNASLIVVLARTEL--PALRYRFNAP----- 792
820 VLHNSVLSLNSQVBPDPDRPPARVAGSWGPCSGSLQKRAVDCRGSGAGORTV-PAC 878
793 -----IARDSLPPYSWHYAPWTKCSAQACAGSGQVQAVECRNQLDSSAVAPHY 839
879 DAHR--PVEQACG-EPC-PTWELSAWSPCSKSGRFFORSILK 920
840 CSAHSLKLPKQACNTEPCPDVWVGNWLSLCSRCDAGVRSRVVC 885

RESULT 12
AT12_HUMAN
ID AT12_HUMAN STANDARD; PRT; 1593 AA.
AC PS8397;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-DE TS12).
DE GN ADAMTS12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=21264577; PubMed=11279086;
RA Cal S., Arqueelles J.M., Fernandez P.L., Lopez-Otin C.;
RT "Identification, characterization, and intracellular processing of ADAM-TS12, a novel human disintegrin with a complex structural organization involving multiple thrombospondin-1 repeats.";
RL J. Biol. Chem. 276:17932-17940(2001).
CC CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
CC CC -!- COPACITOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC CC -!- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely expressed in gastric carcinomas and in cancer cells of diverse origin.
CC CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY SIMILARITY).
CC CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
CC CC -!- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1 DOMAINS.
CC CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC CC -!- SIMILARITY: CONTAINS 8 TSP TYPE-1 DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AJ250725; CAC20419.1; -;
Genew; HGNC:14605; ADAMTS12.

MIM; 606184; -

InterPro; IPR001762; Disintegrin.

InterPro; IPR002870; Pep_M12B_propep.

InterPro; IPR001590; Repolysin.

InterPro; IPR000884; TSP1.

Pfam; PF00090; zpn_MTPeptide.

Pfam; PF01421; Repolysin; 1.

Pfam; PF01562; Pep_M12B_propep; 1.

SMART; SM00209; TSP1; 8.

PROSITE; PS00215; ADAM_MEPRO; 1.

PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

PROSITE; PS00092; TSP1; 2.

PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;

Repeat; Extracellular matrix.

SIGNAL 1 25 POTENTIAL.

PROPEP 26 240 BY SIMILARITY.

CHAIN 241 1593 ADAMTS-12.

DOMAIN 465 544 DISINTEGRIN-LIKE.

DOMAIN 545 596 TSP TYPE-1 1.

DOMAIN 597 700 CYS-RICH.

DOMAIN 701 826 SPACER 1.

DOMAIN 827 881 TSP TYPE-1 2.

DOMAIN 886 943 TSP TYPE-1 3.

DOMAIN 947 995 TSP TYPE-1 4.

DOMAIN 996 1315 SPACER 2.

DOMAIN 1316 1364 TSP TYPE-1 5.

DOMAIN 1367 1423 TSP TYPE-1 6.

DOMAIN 1426 1471 TSP TYPE-1 7.

DOMAIN 1471 1711 TSP TYPE-1 8.

DOMAIN 302 305 POLY-GLU.

SITE 208 208 CYSTEINE SWITCH (POTENTIAL).

METAL 392 392 ZINC (CATALYTIC) (BY SIMILARITY).

ACT_SITE 393 393 BY SIMILARITY.

METAL 396 396 ZINC (CATALYTIC) (BY SIMILARITY).

METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).

CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 125 125 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 215 215 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 485 485 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 685 685 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 790 790 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 951 951 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 1104 1104 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 1275 1275 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 1300 1300 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 1320 1320 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 1371 1371 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 1378 1378 N-LINKED (GLCNAC. .) (POTENTIAL).

CARBOHYD 1503 1503 N-LINKED (GLCNAC. .) (POTENTIAL).

SEQUENCE 1593 AA; 177545 MW; 07FF9F48E63BD83A3 CRC64;

Query Match 24.7%; Score 1275; DB 1; Length 1593;

Best Local Similarity 32.4%; Pred. No. 6.1e-82;

Matches 330; Conservative 155; Mismatches 374; Indels 158; Gaps 37;

QY 21 PEREVVVIRLD-----PDINGRRYWRGPEDSGDGLIFQITAFQEDFYHL 68

Db 50 PEVHVGVKRVASGHFLSYGLHYPTSSRR---KRDLDGSEDWVYIRISHEEKDLFFNL 106

QY 69 TPAQDLAPAFSEHLGVPLQGLTGTSDDLRCFCYSGDVNAEPD--SFAAVSLCGLGRCA 126

Db 107 TVNQGLNSIYNEKRYGNLSHVKNMASSAPLCHLSGTVLQQGTGVGTAAALSACHGLTGF 166

QY 127 FGVRGAEYVISPAPNAPAAQNRSGAHLQRRGVPGSGDPTSRGCVASGWNPAILR 186
Db 167 FQLPHGDFTEPV--KKHPLVEGGYHPHIVRRQKVP--ETKEPT--CGLKDSVNLSQKQ 220
QY 187 ALDPYKPRRAGFGESRRRSRRAKRFVSPRIVETLVVADESWMKVFHAG--LEHYLLTL 245
Db 221 ELWRERHNL--PSRSLR-----RSIKERWVETLVVADTKMIEYHSGENVESYILTI 274
QY 246 LATAARLYRHPISILNPNIVVVKVLLLRDSDGPKVTGNAALTLRNFCWAKKLNKVSFK 305
Db 275 MNMVTGLFHNPSIGNAIHIVVRLILLEEBEGLKIVHHAETLSFCWKQKSNKPSDL 334
QY 306 HPEYDTAILFTKQDLCA--TCTDLGMADVMTGDKPRKSCSVIBDDGLPSAFITAHLE 363
Db 335 NPVHDDVAVLLTRKIDICAGFNPCETGLSLHSGMCQPHRSCNINEDSLPLAFITAHLE 394
QY 364 GHVFNPHDNVK--VCEEVFGKLRANHMSPFTLIQDRANPWSACSAIITDLCSHGDC 422
Db 395 GHSFQIHDGKENDCEPVG---RHPYIMSRQLQYDPTPLTWSKCSSEYITRELDKRGWGC 451
QY 423 LLDQPSKP--ISLPEDLPGASYTLISOQELAFGVGSKPCPYMO--YCTKLWCTKAKGOMVC 480
Db 452 LDDIPAKKGLKSVIAPGVYIVHHQCLQYGNPNTFCQEVNVCOTLMCSVKG----FC 507
QY 481 QTRHFPWADGTSCGEGKCLKGCACVERHNLNK--HRVDGSAWAKWDPYPCSTCGGVOL 538
Db 508 RSKLDAADAGTQCGEKKWCMAKCI---IVGKKPESIPGCGWRGFWPWSHCSTCGAGVCS 564
QY 539 ARROCTNTPANGKYCEGVVRVYKSCNLEPCPSSASGKSFREOCEAFNGYHNSTNRLT 598
Db 565 AERLCNNPEPKFGKYCTGERKRYRLCNVHPCRSEA--PTFRQMCSEFDTVPYK--NEL- 620
QY 599 LAVAVPKYSGVSRDKCKLICRANGTYFYVLAQVVDGTLG--SPDSTSVCVUGCKIK 656
Db 621 --YHFFPIF---NPAHPCELYCRPIDGOFSEKMLDAVIDGTFCEGGNSRNVINGICKM 675
QY 657 AGCDGNLGSKKRFDKCGVCGGDNKCKKVTGLF--TKPMHGYNFVVAIPAGASSIDIRG 715
Db 676 VGCDYEIDSNATEDRCVCLGDGSSCQTVRKFKOKESGYVDIGLIPGARDIRVME--- 733
QY 716 YKGLIGDDNYLAKNSQ--GKYLNGHFVVAVERDLVVKGLSLRLRYSGTGTAVESLQASRP 774
Db 734 ---IEGAGNFLAIRSEDPKYYLNGGFIIQ--WNGNYKLAGTVFYQDRKGD--LEKLMATGP 788
QY 775 ILEPLTVELSVGKMTPPRVRSFYLPKPREDKSSHDPGPGPSVLNSVLSNQVEQ 834
Db 789 TWESVWIOQLLF--QVTNPGIKYEYTIQKD-----GLDNDVEQ 823
QY 835 -----PDDRP-----PA 841
Db 824 MYFWOYGHWTGCTGIRROTACIKKGRGMVKAFTCDPTQPNGRQKCKEKKACPP 883
QY 842 RVVAGSGPCASCG--SGLOKRAVDRCGS--AGORTVPACDAAH--RPVETQACGP--C 894
Db 884 RHWGEWEACATCQPHGEKKRTVLCIQTMVSDEQALPPTDQHLKPKTLKLSNRDILC 943
QY 895 PT-WELSAWSPCSKSCGFGFQFRSLKCVGHGRLRLARQCNLHRKPQLDFCVLKP 950
Db 944 PSDWTVGNWSECVSGCGGVIRSVYTCANHD-----EPCDVTRKPNRSALCGLQOC 995

RESULT 13

AT22_HUMAN

ID AT22_HUMAN STANDARD; PRT; 1211 AA.

AC O95450;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and

DE metalloproteinase with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2)

DE (Procollagen I/II amino-propeptide processing enzyme) (Procollagen I

DE N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (pNPT)

(Procollagen I/II amino-propeptide processing. enzyme).

ADAMTS2 OR PCING OR PCPNI.
Homo sapiens (Human).
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
(1)
SEQUENCE FROM N.A. (ISOFORMS LPNPI AND SPNPI), AND DISEASE.
TISSUE=Skin;
MEDLINE=99347935; PubMed=10417273;
Collige A., Sieron A.L., Li S.-W., Schwarze U., Petty E.,
Wertelecki W., Wilcox W., Krakow D., Cohn D.H., Reardon W.,
Byers P.H., Lapiere C.M., Prockop D.J., Nussgens B.V.;
"Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis
are caused by mutations in the procollagen I N-proteinase gene.";
Am. J. Hum. Genet. 65:308-317(1999).
CC -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior
CC to fibril assembly. Does not act on type III collagen. May also
CC play a role in development that is independent of its role in
CC collagen biosynthesis.
CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC at alpha-1(I) at Pro-1-gln and of alpha-1(II) and alpha-2(I) chains
CC at Ala-1-gln.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBUNIT: MAY BELONG TO A MULTIMERIC COMPLEX. BINDS SPECIFICALLY TO
CC COLLAGEN TYPE XIV (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; LPNPI (SHOWN HERE) AND SPNPI;
CC ARE PRODUCED BY ALTERNATIVE SPLICING. SPNPI HAS NO SIGNIFICANT N-
CC PROCOLLAGEN PEPTIDASE ACTIVITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVEL IN SKIN, BONE, TENDON
CC AND AORTA AND AT LOW LEVELS IN THYMUS AND BRAIN.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- DISEASE: Defects in ADAMTS2 are the cause of Ehlers-Danlos
CC syndrome type VIIC (EDS-VIIC), a recessively inherited connective-
CC tissue disorder characterized clinically by severe skin fragility
CC and joint hypermobility and biochemically by the presence in skin
CC of procollagen incompletely processed at the N-terminus.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)
or send an email to license@isb-sib.ch).

CC EMBL; AJ003125; CAA05880.1; -
CC DR MEROPS; M12.301; -
CC DR Genew; HGNC:218; ADAMTS2.
CC DR MIM; 604539; -
CC DR MIM; 225410; -
CC DR InterPro; IPR001762; Disintegrin.
CC DR InterPro; IPR002870; Pep_M12B_propep.
CC DR InterPro; IPR001590; Reprolysin.
CC DR InterPro; IPR000884; TSP1.
CC DR InterPro; IPR000130; Zn_MTpeptdse.
CC DR Pfam; PF00090; tsp_1; 4.
CC DR Pfam; PF01421; Reprolysin; 1.
CC DR Pfam; PF01562; Pep_M12B_propep; 1.
CC DR SMART; SM00209; TSP1; 4.
CC DR PROSITE; PS50215; ADAM_MEROPS; 1.
CC DR PROSITE; PS50092; TSP1; 1.
CC DR PROSITE; PS00142; ZINC_PROTEASE; 1.
CC DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

KW	Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;	
KW	Repeat; Collagen degradation; Extracellular matrix; Heparin-binding;	
KW	Alternative splicing; Ehlers-Danlos syndrome.	
FT	POTENTIAL.	
FT	SIGNAL 1 29	
FT	PROPEP 30 253	
FT	CHAIN 254 1211	
FT	METAL 408	
FT	ACT_SITE 409 409	
FT	METAL 412 412	
FT	METAL 418 418	
FT	DOMAIN 480 560	
FT	DOMAIN 561 617	
FT	DOMAIN 618 722	
FT	SITE 691 693	
FT	DOMAIN 723 851	
FT	DOMAIN 852 911	
FT	DOMAIN 912 974	
FT	DOMAIN 975 1030	
FT	DOMAIN 40 43	
FT	DOMAIN 185 188	
FT	CARBOHYD 112 112	
FT	CARBOHYD 251 251	
FT	CARBOHYD 949 949	
FT	CARBOHYD 993 993	
FT	CARBOHYD 1031 1031	
FT	CARBOHYD 1098 1098	
FT	CARBOHYD 1145 1145	
FT	CARBOHYD 1150 1150	
FT	VARSPLIC 544 566	
FT	VARSPLIC 567 1211	
FT	SEQUENCE 1211 AA; 134722 MW; BCCEEF25C23CAD2D CRC64;	
QY	Query Match 23.8%; Score 1231; DB 1; Length 1211;	
QY	Best Local Similarity 32.9%; Pred. No. 5.5e-799;	
QY	Matches 345; Conservative 136; Mismatches 399; Indels 168; Gaps	
QY	14 RTAGSGSP-----EREWVPIRLDPD-----INGRR-----YY 41	
QY	38 RLAAADPGCPGLGHERILAVPRTDAQRLVSHVWSAATSRAGVRRARAAVFTPEF 97	
QY	42 WRGPESDGQGLIFQITAFQEDPYLHLTPDAQFLAPAFSTEHLGVLPQLGTLGSSDLRRC 101	
QY	98 PGGNEEPGSHLFVNVTVFGRDLHLRLRPNARLVAPGATMEMOGE--KGTRVPELLGSC 155	
QY	102 FYGSDVN--AEPDSFAVSLCGGLRGAFYGEAYEYISPLPNASAPAAQNSQG-AHLLQ 158	
QY	156 LYVDGAVGLAEAS-VALSNCDDGLAGIRMBEEFFTEPLEKGL--AAQAEQGRVHVY 212	
QY	159 RRGVPGPGSDPTSRGCVASGWNPAILRALDPYKPRAGFGEISRRSRRAKFFVSIPR 218	
QY	213 RRPTSPLGCPQA--LDTG---ASLSDLSLS-RALGVLEEHSNRRRARHAADD 265	
QY	219 Y-VETLVVADESVMKFRGAD-LEHYLTLTLAARLYRHPSILNPINIVVKVLLLRDR 276	
QY	266 YNIEVLLGVDSVVOFHGKEHVQKYLTLNINVIYHDESAGAHNVILVRILLISYKG 325	
QY	277 SGPKV-TGNRAALTRNFCANQKLNKVKSDKHPEWDPAITLFTRODLCGAITCDTLGMADY 335	
QY	326 SMSLIEITGNFSQSLNVCWAYLQOKPDTGDEYHDHAIFLTRDFF-GPGSMQ--GVAPV 382	
QY	336 GTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPNHDVN-KVCEVEFGKLRAHNMSPTL 394	
QY	393 TGMCHPVRVRSCTLINEDEGFSSAFVVAHETGHVGLGNEHDGQGNRCGD---EYRLGSIAPLV 439	
QY	395 IQIDRANPWSACSAIIITDFLDSHGCDLLDQP---SKPTSLPBDLPASYYTILSQQCELA 451	
QY	440 QAAPHRFHWSRCSQOELSRYLHS--YDCLLLDDPPAHDPW-ALPQ-LPGLHYSMNEQCRFD 495	
QY	452 FGVGSKPCPYMQY---CTKLCTCAKAGQWCVQTRHHPWADGTSCGEGKLCCKGACY--- 505	
QY	496 FGLGYMCTAFRTDPCKQLWCS-HPDNPYFCTKGPGLDGTMCAPGKHCFGKHCILWLT 554	


```
QY 506 ----ERHNLKRVDPGSGWAKWDYPCGSRCTCGGQVLARQCTNPTPANGKKYCEGVRYK 561
Db 555 PDILK-----DGSWAWSPGSCSRCTCGTGVKFTQCNDPAMPNGRGTCSGLAYD 606
QY 562 YRSCNLEPCSSAGKSFREOCFAFNGY-----NHSNRTJLAVAWPKYSGVSPRD 614
Db 607 FOLCSRODCPDLSA--DFREOCQWDLYPEHGDAQHH-----WLP-HEHRDAKE 653
QY 615 KCKLICRANGTGYFYLAPKVDGTLGS--PDSTSVCOGKCIKAGCDGNLGSKKRFDKCG 673
Db 654 RCHLYCESRETGEVWSMKRMVHDGTRCSYKDAFSLCVRGDCRVGCDGVIGSSKQEDKCG 713
QY 674 VCGDNKSKVKYGLFTK--PMHGYNEFVAIPAGASSIDIRQRYKGLIGDDNYLALKNS 731
Db 714 VCGDNHCKVWGTFTSRPKKGYIKMFEIPAGARHLLOE-----VDATSHLAVKNL 768
QY 732 Q-GKYLNGHFVVSAYERDLVVKGSLRLRYSGTGTAVESLOASRPILFPTVEVLSVKMT 790
Db 769 ETGKFIINEENDVASSKTFIAMGVWEYRDE--DGRETLQTMGLHCTITVLVIVPGD-T 826
QY 791 PPRVYSFYLPKEP-----RED-----KSSHPKDPRG-----817
Db 827 RVSLTYKMHEDSLNVDNNVLEEDSVYEWALKWSPCRPCGCGSQFTKYGCRRRLD 886
QY 818 PSVLHN---SVLSLSNQVEQ---PDRPPARWVAGSWGPCSASCG--SGLQKRAVDCRGA 870
Db 887 HKMVHRGFCAALSKPAIRACNPQCSQPVVVTGEWPCSTQCGRTGMQVRSVRCIQPL 946
QY 871 GQRTVPACDAAH---RPVETOAGC--EPCP--TWELSAWSPCSKSCGRGQRSLKCVG 924
Db 947 HDNTRSVHAKHNDARPSRRACRELCPGWRAGFWSCSVTCGNGTQERFVPC-----1002
QY 925 GRLLANDQCNL---HRKPQLDFCVLRPC 950
Db 1003 --RTADDSFGICQERPERPARTCRGLGPC 1028

RESULT 14
AT57_HUMAN
ID AT57_HUMAN STANDARD; PRT; 997 AA.
AC Q9URP4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-7 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
with thrombospondin motifs 7) (ADAM-TS 7) (ADAM-TS7).
GN ADAMTS7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP
RX MEDLINE=99395124; PubMed=10464288;
RA Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;
RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, Novel Members of a New Family of
RT Zinc Metalloproteases.";
RL J. Biol. Chem. 274:25555-25563(1999).
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,
CC LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PESTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF140675; AAD56358.1; -
DR HSSP; P15167; IATL.
DR MEROPS; M12.231; -.
DR Gnew; HGNC:223; ADAMTS7.
DR MIM; 605009; -.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR002870; Pep_M12B-propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF01562; Pep_M12B-propep; 1.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS00142; ZINC-PROTEASE; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
DR Repeat; Extracellular matrix.
DR SIGNAL; 1 27 POTENTIAL.
DR PROPEP; 28 232 BY SIMILARITY.
DR FT CHAIN; 233 997 ADAMTS-7.
DR FT SITE; 204 204 CYSTEINE SWITCH (POTENTIAL).
DR FT METAL; 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
DR FT ACT_SITE; 389 389 BY SIMILARITY.
DR FT METAL; 392 392 ZINC (CATALYTIC) (BY SIMILARITY).
DR FT METAL; 398 398 ZINC (CATALYTIC) (BY SIMILARITY).
DR FT DOMAIN; 462 537 DISINTEGRIN-LIKE.
DR FT DOMAIN; 538 594 TSP TYPE-1 1.
DR FT DOMAIN; 595 697 CYS-RICH.
DR FT DOMAIN; 698 914 SPACER.
DR FT DOMAIN; 915 990 TSP TYPE-1 2.
DR FT CARBOHYD; 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
DR FT CARBOHYD; 693 693 N-LINKED (GLCNAC. . .) (POTENTIAL).
DR FT CARBOHYD; 778 778 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 997 AA; 6587044ED02FC104 CRC64;

Query Match 23.58; Score 1215.5; DB 1; Length 997;
Best Local Similarity 35.18; Pred. No. 5.4e-78;
Matches 353; Conservative 117; Mismatches 376; Indels 159; Gaps 46;

QY 1 MLLGLILTAFAAGRTAGGSEPER---EVVVPRLDP-----DINGRR- 39
Db 18 LLLLCALAPGAPGAPGRATEGRAALDIVHPVRVDAGGSLSYELWPRALKKFDVSVRD 77
QY 40 ---YY---WRGPEDSGDGLIFQITAFQEDFYHLTPDAQFLAPAF--STEHLGVPLQGL 91
Db 78 APAEYELQYRGRE-----LRFNLTAQ-----LAPGFVSETRREG---GL 116
QY 92 TGGSSDLR---RCFYSGDVNABPD---SFAVSLCGLRGAGCYRGAEEVISPPLNASA 144
Db 117 --GRAHTRAHTPACHLLGEVQ--DPELEGGAAIASACDGLAGVFLQSNEDYFIPLD--SA 171
QY 145 PAAGRNQGAHLLQRRGVP-----GGPSGDPTRSGCVASGNPAILRALDPYKPRAGFG 199
Db 172 PARPGHAQ--PHVVKQAPERLAQRGDSSAP-STCGV-----QVYPELESRRERW 219
QY 200 ESRSRRSRGRKRF-----VSPRYVETLVVADESQWKFHG--ADLEHYLLTLLATAARLYR 254
Db 220 EQQQRWRPRLRLRHQRSVSKWKVETLVVADAKMVEYHGQPVQSVYLTIMNVAAGLFH 279
QY 255 HPSILNPINIVVKKVLLLRDRSGPKVTGNAALTNRNFCAMQKKNKVSQKHFYEYNDTAI 314
Db 280 DPSIGNPIHITIVRLVLEDEEDLKITHDADNTLKSFKWQKSNKMKGDAPLHHDTAI 339
QY 315 LFTQDLCGATT--CDTLGMADVGTMCDPKRCGSVIEDDGLPSAFTTAHELGHVFNPHD 372
```

Db 340 LLTKDLCAAMNRPCTGLSHVAGMCQPHRSCSINEDTGLPLAFTVAHELGHSGFIQHD 399
Qy 373 -NVKCEVFQKLRANHMSPFLIDRANPWSACSAIITDFLSDCHGCLLDQSKP- 430
Db 400 GSGNDCEPV-GK--RPFIMSPQLLYDAAPLTWRSRCSQYITRFLDKWGLCLDDPPAKDI 456
Qy 431 ISLPEDLPASYSYLSQCELAFGVSGPCPYM-QYCTKLWCTGKAKGMVCQTRHPFPWAD 489
Db 457 IDFPSPVPPGVLYDVSHOCLQYGAISAFCEMDNVCHTLWCSVGT-----TCHSKLDAVD 512
Qy 490 GTSCEGKGLCKLKGACVERHNLKHRVDSWAKWDYPGCSRTGCGVOLARRQCTNTPPA 549
Db 513 GTRCGENKWLSCGECVP-VGFRPEAVDGGWSGSAWSTCSRSGCMGVQSAERQCTQTPK 571
Qy 550 NGGKYCEGVRYKYSCLNLEPCPSASGSKSPREEOCEAFNGYNHSTNRLTLAVAVPKYSG 609
Db 572 YGGRYCVGERKFRCLNLOACP--AGRSFRHVQCSHFDAWYKGLQHT----WPPVVD 625
Qy 610 VSPRDKCLICRANGTYFYVLAQKVDGTLCSPP--DSTSVGVGQKICAGCDGNLGSKK 667
Db 626 VNP---CELHCPANEYAKKLRAVDVDTGTCYQVRASRDLCINGICKNVGCDFEIDSGA 682
Qy 668 RFDKCGVCGGDKCKVGLTFTPMH-GYNFVVAIPAGASSIDIFORGYKGLIGDDNYL 726
Db 683 MEDRCGVCHGNGSTCHTYSGTFFEEAGLGYVDVGLIPAGAREIRIQE-----VAEAAANFL 737
Qy 727 ALKNSQ-GKYLNGHFVYSAVERDLVVKGSLLRYSGTGTAVESLQASRPILPLTVEVL- 784
Db 738 ALRSEDPKIFUNGWTIQ--WNGDQVQAGTFTYARRGN-WENTLSPGTPKEPWIQVPA 795
Qy 785 -----SVGKMTPRVRYSFYLPKPREDEKSSHPKDPGPGSVLHNSVLSNQVEQDD 837
Db 796 SRGPGGSGRGVPRSTLHGRSRPGVSPGSTEPPGPPAAASTSVS----- 845
Qy 838 RPPARK---VA-----GSGWGPCSASCGSLQKRAVDCRS-----AGQRTVPACDAHRP 884
Db 846 -PSLKWPNLVAHVHGGWG--QAPLGLGWRHRLVLMGPRLPTQLLFOESNPGVHYEY-T 901
Qy 885 VETOACGE-----PCTWELSAWSPCSKSCGRG--FORRSILKCVG 922
Db 902 IHREAGGHDEVPPVFWSHYHGWTKCTVTCGKEGWRHSPTCRG 946

RESULT 15
ATS2_BOVIN STANDARD; PRT; 1205 AA.
ID AC P79331;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-2 precursor (PC 3.4.24.14) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2)
DE (Procollagen I/II amino-propeptide processing enzyme) (Procollagen I
DE N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (pNPI).
GN ADAMTS2 OR NPI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=97225960; PubMed=9122202;
RA Collège A., Li S.W., Sieron A.L., Nusgens B.V., Prockop D.J.,
RA Lapiere C.M.;
RT "CDNA cloning and expression of bovine procollagen I N-proteinase: a
RT new member of the superfamily of zinc-metalloproteinases with binding
RT sites for cells and other matrix components.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).
RN [2]
RP PARTIAL SEQUENCE.

RX MEDLINE=95348096; PubMed=7622483;
RA Collège A., Beschlin A., Samyn B., Goebels Y., Van Beeumen J.,
RA Nusgens B.V., Lapiere C.M.;
RT "Characterization and partial amino acid sequencing of a 107-kDa
RT procollagen I N-proteinase purified by affinity chromatography on
RT immobilized type XIV collagen.";
RL J. Biol. Chem. 270:16724-16730(1995).
CC -!- FUNCTION: CLEAVES THE PROPEPTIDES OF TYPE I AND II COLLAGEN PRIOR
CC TO FIBRIL ASSEMBLY. DOES NOT ACT ON TYPE III COLLAGEN. MAY ALSO
CC PLAY A ROLE IN DEVELOPMENT THAT IS INDEPENDANT OF ITS ROLE IN
CC COLLAGEN BIOSYNTHESIS.
CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC alpha-1(I) at Pro-1-Gln and of alpha-1(II) and alpha-2(I) chains
CC at Ala-1-Gln.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SUBUNIT: MAY BELONG TO A MULTIMERIC COMPLEX. BINDS SPECIFICALLY TO
CC COLLAGEN TYPE XIV.
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: ENZYMIC ACTIVITY IS DETECTED AT HIGH LEVEL
CC IN ALL TYPE I COLLAGEN-RICH TISSUES SUCH AS SKIN, BONES, TENDONS
CC AND AORTA AND AT LOW LEVEL IN BRAIN AND THYMUS. THE MRNA LEVELS
CC WERE DISPROPORTIONATELY HIGH IN HEART, LIVER, RETINA AND MUSCLE.
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -!- DISEASE: DEFECTS IN ADAMTS2 ARE THE CAUSE OF DERMATOSPARAXIS, A
CC RECESSIVELY INHERITED DISORDER CHARACTERIZED BY SEVERE SKIN
CC FRAGILITY AND BIOCHEMICALLY BY THE PRESENCE IN SKIN OF PROCOLLAGEN
CC INCOMPLETELY PROCESSED AT THE AMINO TERMINUS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X96389; CAA65253.1; -
CC MEROPS; M12.301; -
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Repolysin.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000130; Zn_Mrpeptase.
DR Pfam: PF000090; tsp_1; 4.
DR Pfam: PF01421; Repolysin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR SMART: SM00209; TSP1; 4.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Collagen degradation; Extracellular matrix; Heparin-binding.
KW SIGNAL 1 28 POTENTIAL.
FT PROPEP 29 253 BY SIMILARITY.
FT CHAIN 254 1205 ADAMTS-2
FT METAL 402 402 ZINC (CATALYTIC) (POTENTIAL).
FT ACT_SITE 403 403 BY SIMILARITY.
FT METAL 406 406 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 474 554 DISINTEGRIN-LIKE.
FT DOMAIN 555 611 TSP TYPE-1 1.
FT DOMAIN 612 716 CYS-RICH.
FT DOMAIN 717 845 SPACER.

```
FT DOMAIN 846 905 TSP TYPE-1 2.
FT DOMAIN 906 968 TSP TYPE-1 3.
FT DOMAIN 969 1024 TSP TYPE-1 4.
FT SITE 685 687 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 31 35 POLY-ALA.
FT DOMAIN 177 180 POLY-GLU.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 942 942 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 943 943 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 987 987 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1092 1092 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1144 1144 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1205 AA; 133887 MW; 7B5B232A45320371 CRC64;

Query Match 23.1%; Score 1192; DB 1; Length 1205;
Best Local Similarity 31.7%; Pred. No. 3.1e-76;
Matches 329; Conservative 134; Mismatches 403; Indels 172; Gaps 42;

Qy 1 MLLGIITLAFAGTAG-----GSEPEREVVPIRLD----- 32
Db LLLLLLLPADARAAAAADPPGPGQGHGAERILAVPVRTDAQRLVSHVSNATAPAG 75
Qy 33 -----PDINGRRYWRGPDSDGDOGLIFQITAFQEDFYHLTPDAQFLAPAFST 81
Db 76 VRTTRAAPAQIPGLSG-----GSEEDPGGRLFYNTVFGDLRLRLRPNARLVAPGATV 129
Qy 82 EHLGVPLQGLTGGSSDLRRRCFYSGDV-NABPDSFAAVSLCGLRGAGFYGAEEVVISPLP 140
Db 130 EQWQE--SGATRVLEPLTCLTVYGDVAGLAESSVALNSCDGLAGLIRMEEEFFIEPLE 187
Qy 141 NASAPAAQRNSQG-AHLLORRGVP-----GGPSGDPTSRGVSAGSNPAIRALDOPYKP 193
Db 188 KGL--AAKEAQGRVHVYHRPTTSRPPPLGGPOALDT---GISADSLSLRAL----- 237
Qy 194 RRAFGFSRSTRSRAKRVSTPRY-VELLVVADESWMVFHGAD-LEHYLLTLLATAAR 251
Db 238 ---GVLEERVSSRRRRRAADDDYNIIEVLGVDDSVQFHGTEHVQKYLTLTMNIVNE 294
Qy 252 LYRHPSTLNPINIVVVLRLORDSGPKV-TGNAALTLENFCAWOKKLNKVSDDKHPYEW 310
Db 295 IYHDESIGAHINIVVVLRIILSVKSMLEIGNPSOSLENVCRWAYLOQKPDPTDHEYH 354
Qy 311 DTAILTRQDLGGATTCDTLGMADVGTMCDPKRSVIEDDGLPSAFTTAHELGHVFNMP 370
Db 355 DHAIFLTRQDF-GPSGMQ--GYAPVTGCHPVRSCTILNHEDGFSSAFVVAHETGHVLMGE 411
Qy 371 HDNV-KVCEEVFGKLRANHHMSPTLQIDRANPWSACSAIITDFLDSHGDCLLDQP-- 427
Db 412 HDGQGNRCGD---EVRIGSIAPLVQAAAFHFHWSRCSQOELSRYLHS--YDCLRDDPFT 466
Qy 428 -SKPISLPEDLPGASVYLSQOCELAGVSGKPCPYMQY---CTKLNCTGKAKQMVQCTR 483
Db 467 HDWP-ALPQ-LPGLHYSNNQCRDFGLGYMMCTAFRTDPCKQLWCS-HPDNPYFCKTK 523
Qy 484 HFPWADTSCGEGKCLKLGACV-----ERHNLNKHVRVDGWSNAKWDPYPCSRCTCGGV 536
Db 524 KGPPLDGTMCAPGKHCFCGHCILWTPDLKR-----DGNWGAWSFGSCSRTCTGTG 575
Qy 537 QLARRQCTNTPANGKYCGVRVRYKRSNLEPCPSASGKSFPREECEAFNGY----- 590
Db 576 KFRTRQCDNPHANGGRTCSGLAYDQLCNSQDCPDALA--DFREECRCQWDLYFEHGDA 633
Qy 591 -NHSTNRLTLAVAVPKYSGVSPRDKCLICRANGTGYFYVLAPKVVDGTLCG-PDSTSV 648
Db 634 QHH-----WLP-HEHRDAKERCHLYCESKETCEVVVSMKRMVHDGTRCSYKDAFSL 682
Qy 649 CVQCKICKAGCDGNLGSKKRFDKCGVCGGDNKSKCKVTGLFTKPMH--GYNFWVAIPAGA 706
Db 683 CVRGDCRKCVDGVISSKQEDKCGVCGGDNHCKVVKVGTFSRSPKGLGYIKMFEIPAGA 742
```

```
Qy 707 SSTDIRQRYKGLIGDNDYLALKNQSG-KYLLNGHFVWSAVERDLVVKGSLRLRYSGTGA 765
Db 743 RHLLIQEAD-----TTSHLAVKNLETKFILNEENDVDPNKSTFTTAMGVWEYRDE-DG 796
Qy 766 VESLOASRPILLEPLTVEVLVSGKMTPPRYRYSFYLPKPEPREDKSSHPKDPGPGSVLHNSV 825
Db 797 RETLQTMGPLHGTITVLVIPEG---DARISLTY-----KYMIHEDS 834
Qy 826 LSLSNQVEQPDPPPARVWAGSWGPGCSAGSGLOKRAVDRCGSAGQRTVPA--CDAAH 883
Db 835 LNVDDNNVLEDDSVGVENALKWSPCKSCGGGQFTKYGCRRRLDHKMWHRGFCDSVSK 894
Qy 884 P-----VETOAGCEPCTWELSAWSPCKSCGR-GFQRRSLKCV-----GHGGRLLARDQ 932
Db 895 PRAIRRTCNPQECQ--PVMVTGEWEPESRSCGRTGMQVRSVRCVOPLHNNTTRSVHTKH 952
Qy 933 CNLHRKPQELDFCVLRPC 950
Db 953 CN-DARPEGRACNREL 969
```

Search completed: May 9, 2003, 15:22:42
Job time : 22 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:20:56 ; Search time 27 Seconds
(without alignments)
3382.507 Million cell updates/sec

Title: US-09-965-631-4
Perfect score: 5162
Sequence: 1 MLLGILTFLAFAGTAGGSE.....DQNLHKKPOELDFCVLRPC 950
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2485.5	48.1	951	2 T00017	gene ADAMTS-1 prot
2	1913	37.1	837	2 T00355	hypothetical prote
3	1604	31.1	550	2 T47158	hypothetical prote
4	1587	30.7	2165	2 T21371	hypothetical prote
5	1192	23.1	1205	2 T18517	procollagen N-endo
6	681	13.2	1558	2 C89114	protein C37C3.6a
7	681	13.2	2167	2 T34395	hypothetical prote
8	559.5	10.8	1444	2 T18856	angiogenesis inhib
9	510.5	9.9	860	2 T18892	hypothetical prote
10	493	9.6	951	2 T00260	hypothetical prote
11	450.5	8.7	957	2 T15976	hypothetical prote
12	411.5	8.0	1059	2 T22545	hypothetical prote
13	343.5	6.7	571	2 S24789	Jararhagin C precu
14	340.5	6.6	609	2 S55270	catrocollastatin p
15	324	6.3	903	2 S60257	maltin alpha - mo
16	323.5	6.3	617	2 S48160	metalloproteinase
17	320	6.2	826	2 A60385	monocyte surface a
18	313	6.1	549	2 S48169	metalloproteinase
19	311	6.0	411	1 HYSNFA	fibrolase (EC 3.4.
20	308	6.0	407	2 S66260	metalloproteinase
21	307	5.9	789	2 S28259	androgen-regulated
22	304.5	5.9	616	2 A55796	ecarin precursor -
23	298.5	5.8	480	1 A30065	trigramin precursor
24	292	5.7	478	2 JC4880	fibrinolytic metal
25	291	5.6	414	2 S41609	atrolysin C (EC 3.
26	290.5	5.6	481	2 JC4342	fibrinolytic prote
27	289.5	5.6	610	2 JC7530	vascular apoptosis
28	286	5.5	414	1 HYSAC	atrolysin C (EC 3.
29	284	5.5	481	2 S43125	trimucin precursor

30	277.5	5.4	814	2 G02390	disintegrin-like m
31	277	5.4	478	2 A43296	atrolysin E (EC 3.
32	267	5.2	414	2 S41608	atrolysin B (EC 3.
33	265	5.1	670	2 I65967	disintegrin-like m
34	256.5	5.0	1074	2 JC5928	semaphorin F precu
35	256	5.0	776	2 S28258	androgen-regulated
36	255.5	4.9	1584	2 T00026	brain-specific ang
37	253.5	4.9	952	2 T18900	disintegrin and me
38	251	4.9	1042	2 T26644	hypothetical prote
39	248.5	4.8	1170	1 TSHUP1	thrombospondin 1 p
40	247	4.8	478	2 JQ1301	hemorrhagic protel
41	245.5	4.8	1170	2 A40558	thrombospondin 1 p
42	244.5	4.7	419	2 A59414	metalloproteinase
43	236.5	4.6	823	2 S18968	cytostatin precurs
44	234.5	4.5	508	2 T22836	hypothetical prote
45	234	4.5	419	2 S41607	atrolysin A (EC 3.

ALIGNMENTS

RESULT 1

T00017
gene ADAMTS-1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00017
R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.
Genomics 46, 466-471, 1997
A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1
A:Reference number: Z14055; MUID:98110583; PMID:9441751
A:Accession: T00017
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-951 <KUN>
A:Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057
A:Experimental source: strain 129SVJ
C:Genetics:
A:Gene: ADAMTS-1
A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2
C:Superfamily: thrombospondin type 1 repeat homology
F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match	48.1%	Score	2485.5;	DB 2;	Length	951;			
Best Local Similarity	48.0%	Pred. No.	2.7e-161;						
Matches	478;	Conservative	159;	Mismatches	248;	Indels	111;	Gaps	22;
QY	1	MLLIGILTLAGAGTAGG--SEPERVVVPIRLDPDINGRRYYWRCP-EDSGDQGLIFQI	57						
Db	20	LLLLASITMLLCARGAHRPTDEEELVLP-SLE-----RAPGHDSITTRL--RL	66						
QY	58	TAFQDEFYHLTPDAQFLAPAFSTEHGLGVPLQGLTGS-----SDLRRRCFYSGDVNA	109						
Db	67	DATGQQLHLKLPDGSGLFAGFTLTQV-----GRSPGEAQHLDPDGLAHCFYSGTVNG	121						
QY	110	EPDSFAAVSLCGLRGAFGRGAEYVISPUNAS-----APAAQRNSQGA-----HLQRRG	161						
Db	122	DPGSAALSCEGVRGAFYLGQEEFFIQAPGVATERLAPAVPEEESARPQPHILRRR-	180						
QY	162	VPGGPSGDPTRSGCVASGMNPAILRALDPYKPRRAGFGESRSRRR-----	206						
Db	181	-----RRSGGAKCGVMD-----DETLPISDSRPESQNTNRNQPVRDPTPDAGKP	226						
QY	207	-----SGRAKRFVSIPIRYVETLVVADESWMKFGAGADLEHYLLTLLATAARLYRHPSTLNPI	262						
Db	227	SGPGSIRKKRFVSSPRIVETMLVADQSMADFHGSGLKHLYLLTFLSVAARFYKHPSTIRNSI	286						
QY	263	NIWVKVLLLRDSDGPKVTGNAALTNRNFCWAKKLNKVSOKHPYWDNTALFTRQDLG	322						
Db	287	SLVVVILVIYEOKGPVETSNAAFLRNFNCWQKHNPSDRDPDHYDTAILFTRQDLG	346						
QY	323	GATTCDTLGMADVGTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMHPDNVKKVCEEVFG	382						

```
Db 347 GSHCTDTLGMADVGVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCASLING 406
QY 383 KLRANHMSPTLIQIDRANPWSACSAAIITDLDGSHGDCLLDQPSKIPSLPEDLPFGASY 442
Db 407 VTGDSHLMASMLSSLDHSQWSPSCSAIYMYTSLDNGHGECLMDKPNQPIKLPSPDGLTY 466
QY 443 TJSQCELAAGVGSRCP- YMOYCYTKLWCTGAKAGOMYQCTRHFPWADGTSCEGEGKCLK 501
Db 467 DANROQFTFGESKCHPCDAATCTTLWCTGTSGGLLVQCTKHFPWADGTSCEGEGKCVS 526
QY 502 GACVERHNK- --RVDSWAKWDPYPCSRRTCGGQVQARROCTNPTBANGKCYCEGV 558
Db 527 GKCVNKTDM- KHATPVHSGWGPWGDGCSRTCGGVOYTMRECDNPVENKNGKCYCEGK 585
QY 559 RVKYSNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVARVPKYSVGSVPRDKCKL 618
Db 586 RVYRSCNIEDCPDN- NGKTFREEQCEAHNEFSKASFGNEPTVENTPKYAGVSPKDRCKL 644
QY 619 ICRANTGTFFYLAPKVVDTLCSDDSTSVGVGKCIKAGCDNGLGSKKFPDKCVCGGD 678
Db 645 TCEAKGIGYFFLQPKVVDGTFCSDDSTSVGVGQCVKAGCDRIIDSKKFKDKCVCGGN 704
QY 679 NKCKKVTGLTKPMHGFNFVAIPAGASSIDIROGVKGLIGDDNYLALNKSQCKYLLN 738
Db 705 GSTCKKMSGIVSTRPGYHDIVITPAGATNIEVKHNRQGSERNNGSFLAIRAADSTYILN 764
QY 739 GHFVSAVERDLVKGSLRYSGTGTAVESLQASRIILEPLFVLSVGMKMTPPRVYSF 798
Db 765 GNFTLSTLEQDLTKYKTVLRYSGSSAALIRISFSPKLEPLTIQVLMVGHALRPKIKETY 824
QY 799 YLPKPREDKSHHPDRGFSVLHNSVLSNQVEQDPPARVAGSWGSCSACSGS 858
Db 825 FMKKKTES-----FNAIPTFS-----FWIEEWGECSTKCGSG 857
QY 859 LQRRADVRCGSAQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSGRGFOR 915
Db 858 WQREVVOQRDINGH---PASECAKEVKYPASTRPCADLPCPHQVGDWSPCSTCKGKYK 914
QY 916 RSLKCVGHGGRLLARDQCNLHRKPQ- LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSNESCDPLKPKRHVIDFCTLTQC 950

RESULT 2
T00355
hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, N.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
C:Accession: T00355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <LSH>
A:Cross-references: EMBL:AB014588; NID:g3327189; PIDN:BAA31663.1; PID:g3327190
A:Experimental source: brain
C:Genetics:
A:Gene: KIAA0688
C:Superfamily: thrombospondin type 1 repeat homology
F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 37.1%; Score 1913; DB 2; Length 837;
Best Local Similarity 46.3%; Pred. No. 2.3e-122;
Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

QY 1 MLLGLITLTAFAAGTGGSEPEREVVPIRLDPDINGRRYWRGPDSDGQGLIFQITAF 60
Db 37 LLLLLASLPSARLASPLPREIEIVPEKLSVLT-----PGSGTFAIRLLCRLQAF 88
QY 61 QEDFYLLHLPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRRCFYSGDVNAEPDSFAVSLC 120
```

```
Db 89 GETLLLEQDSGVQVEGLTVQYLGAPE--LLGAEP--GTYLTGTINGDPESVASLHWD 145
QY 121 GG-LRGAFGYRGAEYVTSPLPNASAPAAQRNSOAHLLQRRGVPGSPGSDPTSRCGVASG 179
Db 146 GGALLGVLYRGAEHLHQLPLEGGTPNSA--GGPCAHLIRK-----SPASQGGPCNV--- 196
QY 180 WNPAILRALDPYKPRRAGFGESESRRRSGRAKRFVSIPIRYVETLVVADESWMVHFAGDLE 239
Db 197 -----KAPLGSPPRPR--RAKRFASLSREVETLVVADDKMAAFHAGL 239
QY 240 HYLLTILATAARLYRHPSILNPINIVVKKVLLLRDRDSGPKVTGNAAALTNRNCAWOKKL 299
Db 240 RYLLTVMAAAKAFKHPISIRNPSVLSVLRVLVILSGEGEPQGVGSAQTURSCFAWORG 299
QY 300 NKVSDKPEYWDTAILFTRODLGCAITCDTLGMADVTGMDPKRSCSVIEDDGLPSAFIT 359
Db 300 NTPEDSDPHDFDAILFTRODLGCVSTCDTLGMADVTGCDPARSCAIVDDGLOSFA 359
QY 360 AHELGHVFNPHDNVKEEVFKL-RANHMSPTLIQIDRANPWSACSAIITDLDG 418
Db 360 AHELGHVFNHLDNSKPCISLNGPLSTSRHVMAFVMAHVDPPEPWPSPCSARFITDLDNG 419
QY 419 HGDCLLDQPSKIPSLPEDLPASVTLSSQCELAFGVSKPCPYM-OYCTKLWCTGRAKQ 477
Db 420 YGHCLLDKPEAPLHLPTVTFGKYDADRCQLTFGPDPSRHCPLPPPCAALWCSGHLNGH 479
QY 478 MVCOTRHPWADGTCGEGKCLKGACVERHNLNKHRVD--GSWAKWDYGPSCSRTCGG 535
Db 480 AMCQTKHSPWADGTPCGPAQACMGRCGLHMDQLQDENIPQAGGWPWGWDGDCSRTCGG 539
QY 536 VOLARROCTNPTPANGKYCEGVYKYSNCLBPCPSASGSKSFREEQCEAFNGYNHSTN 595
Db 540 VQFSRDCRTPVPNGGKYCEGRTFRSCNTEDCP-TGSALTFREEQCAA---YNHRTD 595
QY 596 RLTL---AVAWPKYSGVSPDKCKLICRANGTYFYVLAPKVVGDGTLCSPDSTSVCVQ 652
Db 596 LFKSFGPMDWVPRTGVAPQDCKLTQARALGYIYVLEPRVVDGTPCSPDSSVVCVQ 655
QY 653 KCIKAGCDNGLGSKKRFDKCGVGDNKSCKVYGLTKPMHGVNFVVAIPAGASSIDIR 712
Db 656 RCIHAGCDRIIGSKKFKDKCMVCGDGGSGCSKQSGSFRKFRYGYNNVVTIPAGATHILVR 715
QY 713 QRGVYKGLIGDDNYLALNKSQCKYLLNGHFVVSVAVERDLVVKSL-LRYSCTGTAVESLQA 771
Db 716 QQGNPG--HRSIYLALPLDGSYALNGEYTLMPSPDVLVPGAVSLRYSGATRASETLSG 773
QY 772 SRPILPTLTVELSVGKMTPPRVRYSFYLPKPREDKSSHPKD 814
Db 774 HGPLAQPLTLQVLVAGNPQDTRLRYSFVPRPTSTPRTPQD 816

RESULT 3
T47158
hypothetical protein DKFzp762C1110.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47158
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: 224379
A:Accession: T47158
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-550 <AAA>
A:Cross-references: EMBL:AL162080
A:Experimental source: adult melanoma (MeWo cell line); clone DKFzp762C1110
C:Genetics:
A:Note: DKFzp762C1110.1

Query Match 31.1%; Score 1604; DB 2; Length 550;
Best Local Similarity 49.7%; Pred. No. 1.6e-101;
Matches 284; Conservative 106; Mismatches 142; Indels 40; Gaps 11;
```


A;Molecule type: mRNA
A;Residues: 1-951 <NAG>
A;Cross-references: EMBL:AB011177; NID:g3043733; PIDD:BA25531.1; PID:g3043734
A;Experimental source: brain
C;Genetics:
C;Note: KIAA0605
C;Superfamily: thrombospondin type 1 repeat homology
F;46-106/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 9.6%; Score 493; DB 2; Length 951;
Best Local Similarity 22.2%; Pred. No. 1.5e-25;
Matches 171; Conservative 55; Mismatches 183; Indels 360; Gaps 27;

QY	517	GSNAKWDPGPCSRCTGGGVLARRQC----	TNP	PANGKYCEGVRRVYKSCNLEPCPS	572
Db	51	GEWTKWTAF--SRS CGGVT	SQRHCL	QQRRKSYPGCGNRCTGTSKR YOLCHVQCP-	106
QY	573	SASGKSFR EOCFAFNG--YHNSTNRLTLAVAPVKYSVGSPRD-----	KKLCICRAN	623	
Db	107	-PDGRSFREBOCVFSFNHVNGRTHO-----WKPLY-----	PDDVVIHSKKPCDLHC-TT	154	
QY	624	GTCGYFVLAKPVWDGLCS-PDSTSVCVOGCKIKAGCDNLGSKRKREDKGCVGCDNKSC	682		
Db	155	VDOGROLMP-ARDGTSCKLTDLRGVCSGKECPTGCDGVLFSTHTLDKCICOGDGSSC	213		
QY	683	KVVTGLFK-PWH-GYNFVAIPAGASSIDIIRQRYKGLIGDDNYLALKNSQOKYLNGH	740		
Db	214	THVTGNRYKGAHLGYSLVTHIPAGARDIQIVERAKSA-----DV LALADEAGYTFENG	268		
QY	741	FVVS AVERDLVKVSGLLRYSST---GTAVESILQASRP TLBPLTVFLSVL-S	GMWTPPRVR	795	
Db	269	YKVDS-PKNFNIA GTVVYKRRPMDVETGIEYIVAQGPTNOGLNVWMNONGK--SPSIT	325		
QY	796	YSFYLPKPREDKS-----	809		
Db	326	FEYTLLQPHESSRPOIYYGFSSAESQGLDGAGLMGFPHNGSYLQASERLGLDNRL	385		
QY	810	-SHP-----	812		
Db	386	FGHPGLDMELPGSOGETNEVC EQAGGACEGEP RPKGFRDNRNVTGTLTGDKDDBEVDI	445		
QY	813	--KO-----	PR-----	816	
Db	446	HFASOEFSANAISDLLGAGSLDKFTLN TVNSIFAQAPRSSLAESFFV DYERNEGA	505		
QY	817	GPSVLHNSVLSS-----	NQVEQDDRPRA-----	841	
Db	506	GPYLLNGSYLELSDRVANSSEAPPNVYSTLSLT SAGR NTHKART PKARKOGVSPADM	565		
QY	842	-----	-----R	842	
Db	566	YRWKLSSHEPCSATCTTGVM SAYAMCVRYDGVEDVDSDCYDALTRPEVPVHFCAGREQCR	625		
QY	843	YWAGSWGPC SASCGSLGKRAV-----	864		
Db	626	WETSWSCSRTCGEYQFVRVWCWMLSPGDFS SVYSDLCEAAEV RP ERKTCRN PAC	685		
QY	865	-----DCRGSAGRTV-----	PACDAARHPVETQAQ--GEPPCP-TWELSA	901	
Db	686	GPQWEMSEWSECTACKGERSVVRTDRICSEDEKLDPNTRPVGEKNKTGPPCDROVTUSD	745		
QY	902	NSPCKSGRCGRFORSLKCVGHGGRLLARDQNLRHKQEQLDFCVLRPC	950		
Db	746	NGPCSCGCGORTLRHVYCKTSDGRVVPBSOCQMETKPLAHPGCGDKNC	794		

RESULT 11
T15976
hypothetical protein F08C6.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15976
R:Bentley, D.

submitted to the EMBL Data Library, June 1995
A;Description: The sequence of *C. elegans* cosmid F08C6.
A;Reference number: Z18440
A;Accession: T15976
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-957 <BEN>
A;Cross-references: EMBL:U29378; NID:g868184; PID:g868185; PIDN:AAA68721.1; CESP:F08C6
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:F08C6.1
A;Introns: 23/1. 135/3. 285/3. 325/2. 376/3. 451/2. 488/3. 532/3. 527/3. 714/1

Query Match	8.7%	Score 450.5	DB 2	Length 957
Best Local Similarity	22.7%	Pred. No. 1.2e-22		
Matches 227	Conservative 106	Mismatches 328	Indels 339	Gaps 47
QY	57	ITAFQEDFYLHLTPDAQFLAPAFSTEHLGVPLQGLTGSSD----	LRRCFYSGDVNAEPD	112
Db	61	IQAFNKYNLSL----	EXTLAKLSSGVVYVKKNEKGGSLDFSTLDSCHYH----	HYCEK 114
QY	113	SFRAVSLCGGL-----	RGAFYGRGAEYVISPUNASAPAAQRNSQ-GAHLQRRG	161
Db	115	VYAISNCDRIVSESNRKQKGTVIDDGGIIVVHPDPDHAHRSKRATENGAAHVYKRE	174	
QY	162	VPGGSGDPTSRGCVASGWNPAILRALDPVKPRRAGFGESSRRSRGRAKRFVSIPIR----	218	
Db	175	T---LAGEPKDFCGLDN-----	VVTEBSLVEDESAIFEDV-----	FVQGRLTQ 215
QY	219	---YVETLVADESVMKF-----	HG--AD---LEHYLLTLTLAARLYRHPHSILPINIVV	266
Db	216	QSDLIVELAVFVDENLWRHFSKXHGGMADRKLQDYTLTLNNIQIMYQPTASPLTFRV	275	
QY	267	VKVLILLDRDSG---PKVTGNAALTLRNCANQKLNKYSDKHPEYWDTAIFTRODL-	321	
Db	276	IRYEVLTRQPSALAGYLHNNHGNAGMYLDRFCRYQRLN-AVRD-----	WDHAIMLTGYDIH	329
QY	322	CGATTCDTLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFMNPHNDKVCVEVF	381	
Db	330	RGAGRSRSISGIARLDGCMCDPWNTCTLAEGLDFTSAFIGTHELGHRFTELKDH-----	381	
QY	382	GKLRANHMSPTLIQIDRANPWSACSAIITDFLDSGHGCDLQDQSKPISLSEDLPGAS	441	
Db	382	-----KSDTLF-----	-----	387
QY	442	YTLSSQCELAFGVGSKPCPYMQYCTKLWCTKAKGMVQCOTRHPFWADGTSCEGKLCCK	501	
Db	388	-----	TGTFCEGSKWCQL	400
QY	502	GACVERHNLNK-----	HRVDGSKAKWDPYGPCSR-TCGG-----	GWQLA 539
Db	401	GRCPVPTGNEIOPTVQHVAPVVTTLPSRIDGSSWGWA-TICSQCTNGILGSVGLAIA	459	
QY	540	RROCTNPTTANGKCYCEGVKRYKSCNLEPCPSASGSKSFREQC-----	EAFNGY	590
Db	460	RRTCSAPYPANGGDCVGTSTRVILCSRQCGRAKSKSVDHEYISDKMEQKRLKNDRLETKG	519	
QY	591	NHSTNRLTLAVAWPKYSGVSPRDKCKLTCRAN---GTGYFY-VLAPKVVWDGTLCSPDST	646	
Db	520	GSOLNRF-----	PORACKVFCVQOYHGSORNYRFFGDNLPDGTSCGYD--	563
QY	647	SVCVQCKCIKAGCDGNL-----	GSKKRFDC-GVCGGD	678
Db	564	RYCLDGECLALNCNNALISRQSCPTDCTPTDQSSSVYRGOWGTWSLWTSCTATCGG	623	
QY	679	----NKSKCKVTGLFTKPMHGNFVVAIPAGASSIDIQRGYKGLIGDNNYLALKNKSGK	734	
Db	624	YRKRNRAC-SITG---QCEGNEDETEVCSSPCSPV-----	LRVGNWST	664
QY	735	YLLNGHFVVSAYERDLVVGKSLRLYSGTGTAVESLQASRPILP-----	LTVEVLSVGKMTPT	791
Db	665	WTEWNHCYSVC-----GRGQARY-----	RKCLSPHRTLAFDPCGENKVT-704	

QY 792 PVRYSFYLPKPRDKSKHPKDPGRSPSVLHNSVLSLNOVEOPIDRRP-----ARWVA 845
Db 705 NELRIFF-----KARSYIMCSVRCKIKRNTISEKKIEVRSCLNGPCNAIGVWGTV-- 756
QY 846 GSWGPCSASCSSG--LQKRAVD---CRSGAGORTVPACDAARHPVETQACGEPCT---W 897
Db 757 GCMSTCSTSCGPGTLVRQTCNRPCDGSABHRR--SCNVA-----TCNDGIGW 803
QY 898 EL-SAWSPCSKSGRGFORSLKVCVGHG--GELLARDQCN 934
Db 804 SLWNWSDCSRVCGGLRSRSRSCFGSGOMGASSEQFCN 843

RESULT 12
T22545
hypothetical protein F5386.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22545
R:White, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19578
A:Accession: T22545
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1059 <WIL>
A:Cross-references: EMBL:Z81086; PIDN:CAB03121.1; GSPDB:GNO00019; CESP:F5386.2
A:Experimental source: clone F5386
C:Genetics:
A:Gene: CESP.F5386.2
A:Map position: 1
A:Introns: 38/3; 92/3; 131/3; 169/1; 236/2; 282/3; 349/2; 376/2; 420/1; 576/3; 759/3; 843/3
Query Match 8.0%; Score 411.5; DB 2; Length 1059;
Best Local Similarity 25.7%; Pred No. 6.1e-20;
Matches 112; Conservative 53; Mismatches 128; Indels 143; Gaps 16;

QY 518 SWAKWDPIGPCSRCTCGGVQLARRQCTNTPANGKYCEGVRYKRVSCNLEPCPSSASGK 577
Db 25 SWAASPSWSSCTKTCGGVSRQLRRCLT-----SKCSGESVREKVCQAQKTCESKS--R 75
QY 578 SFREEOCEAFNGYHSTNRLTLAVAWPKYSGVSPRDKCLCRANGCYFYVVLAPKYVD 637
Db 76 LARDTICGG-----EEIVSRGQCEVVCVRSLKLGANFLW--RVDD 112
QY 638 GTLC-SPDSTSVQVQCKIKAGCDNLGSKRPFDKGCGVCGDNKCKKVTGLFTRPMHGY 696
Db 113 GTPQATSPAVCSKSCQIVGCDGLISSERFEDAGCVCGGRGDIC----- 158
QY 697 NFVVAIPAGASSIDIRQGYKGLIGDNYLALKNSOGKYL--LNGHFVVVSVAVERDLVYKG 754
Db 159 -----DN-----GKFIVKVSSEYTTACASNCDDIVD- 183
QY 755 SLLRYSGTGAVESLQASRPILPLVEVLSV-GKMTTPRVRYSFVLPKPREDKSSHPK 813
Db 184 -----WSGAGNSIATS-----QPIVVCVNAITGRVVPKELCADKLRPK----- 222
QY 814 DPRGSPVLHNSVLSLNOVEQDDRP-----PARWAGSWGPCSCSGSLQKRAVDGR 867
Db 223 -----VEARPCPMLICPSRMAADWTECPHCCEGTRKREVYCV 261
QY 868 GSAGORTVPA---CDAARHPVETQAC-GEPCPTWLSAWSPCSKSCGGRGFORSLKCVG 922
Db 262 QTAHNVTVHVPDPTFCNGTRPAAEENCVSTSGRWEAGKWSKCTASCQGVGVRHRRHACVG 321
QY 923 -----HGGRLLRDQC 933
Db 322 GSDCDEGRPRQETTC 337

RESULT 13
S24789
jararagin C precursor - jararaca (fragment)

N:Alternate names: single chain botrocetin
N:Contents: disintegrin-like 28K protein; hemorrhagic proteinase (EC 3.4.24.-)
C:Species: Bothrops jararaca (jararaca)
C:Date: 20-Feb-1995 #sequence_revision 29-Aug-1997 #text_change 09-Jun-2000
C:Accession: S24789; J02245; A44463; A37958; J02373
R:Paine, M.J.I.
submitted to the EMBL Data Library, August 1992
A:Reference number: S24789
A:Accession: S24789
A:Molecule type: mRNA
A:Residues: 1-571 <PAI>
A:Cross-references: EMBL:X68251; NID:g62467; PID:g62468
R:Usami, Y.; Fujimura, Y.; Miura, S.; Shima, H.; Yoshida, E.; Yoshioka, A.; Hirano, K
Biochem. Biophys. Res. Commun. 201, 331-339, 1994
A:Title: A 28 kDa protein with disintegrin-like structure (jararagin-C) purified from
A:Reference number: J02245; MUID:94256999; PMID:8198592
A:Accession: J02245
A:Molecule type: protein
A:Residues: 360-571 <USA>
A:Experimental source: venom
R:Paine, M.J.; Desmond, H.P.; Theakston, R.D.; Crampton, J.M.
J. Biol. Chem. 267, 22869-22876, 1992
A:Title: Purification, cloning, and molecular characterization of a high molecular weight
ily.
A:Reference number: A44463; MUID:93054601; PMID:1385408
A:Accession: A44463
A:Molecule type: mRNA
A:Residues: 1-23,'Q',94-131,'G',133-169,'Q',171-571 <PA2>
A:Cross-references: GB:X68251
A:Experimental source: venom gland
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBI:118104)
R:Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.;
Biochemistry 30, 1957-1964, 1991
A:Title: Isolation and chemical characterization of two structurally and functionally
A:Reference number: A37958; MUID:91129280; PMID:1993206
A:Accession: A37958
A:Molecule type: protein
A:Residues: 360-372,'E',374-378,'X',380-384 <FUJ>
A:Note: 361-Val was also found
C:Comment: Inhibits collagen- and ADP-induced platelet aggregation.
C:Superfamily: mouse meltrin alpha; disintegrin homology
C:Keywords: hydrolase; metalloproteinase; venom; zinc
F:360-571/product: jararagin C #status experimental <MAT>
F:362-444/Domain: disintegrin homology <DIS>
F:295,299,305/Binding site: zinc (His) #status predicted
F:296/Active site: Glu #status predicted

Query Match 6.7%; Score 343.5; DB 2; Length 571;
Best Local Similarity 22.7%; Pred No. 1.2e-15;
Matches 160; Conservative 88; Mismatches 256; Indels 201; Gaps 35;

QY 12 AGRTAGSEPEREVVPIRLDPDINGRRYWRGPDGSDQGLIFQITAFQEDF-YLHLTP 70
Db 1 ATRPKGAVQPKYEDA--MOYEFKVNQEPVVLHLEKNKG-----LFSKDYSEIHVSP 49
QY 71 DAQFLAPAFSTHEGLVPLQGLTGSSDLRCFYSDVNAEPDSPAAYSLGGLRGAFGYR 130
Db 50 DGEITTYPPVED-----HCYHGRLENDADSTASISACNGLGYEKLQ 93
QY 131 GAETVISP--LPNAPAAQNSQGAHLLQRRGVPGGSDPTSRGCVASGN----- 181
Db 94 RETFIEPLKLPDSEAH-----VFKYENVE--KEDEAPKMGCVTQNKSYRIFIKK 142
QY 182 -----PAIRALDPYPRRRAGFSGESRRRSRAKRFVSIPIRVETLVVADESNNK-HG 235
Db 143 ASQLAFTAEQQRDPYK-----YIEFFVVDQGTVTNN 176
QY 236 ADLEHYLLTLATARLYRHPISILNPI-NIVVVKVLL--LRDRSGPKVT--GNAALTLR 290
Db 177 GDLD-----KIKARMYELANIVNEIFRYLMHVALVGLIWSNGDKIIVKPDVDTLN 229
QY 291 NFCAMQKLNKVSOKHPEYWDTAILFTQDLGCGATTCTDLGMADVGTMCDFKSCSVIED 350

Db 230 SFAEWRKTDLETRKKH-----DNAQLLTAIDFNG-----PTIGYAYIGSMCHPKRSVGIVQD 281
QY 351 ---DGLPSAFTTAHELGHVFNMPHDNVKVEEVEFKLRA NH---MMSPTLIQIDRANPWS 404
Db 282 YSPINLVAVVIMAHMGHNLGIIHDT-----GSCSDGYDCIMGPT- ISNEPSKFFS 332
QY 405 ACSAAITDFLDSGHGDCLLDOP-----SKPISLPEDL-GASYTLSSQOCELAFGVGSKP 458
Db 333 NCSYIQWDFIMNHNPECILNEPLGTDIISPPVCGNELL-----EVEECD----- 378
QY 459 CPYQWYCTKLWCTCKAKQWVCQTRHPFPWADGTSCGGKLCCLKGACVERHNLAHRVD-- 516
Db 379 CGTPENCQNECCDA-----ATCKLK-----SGSQCGHGD-----CCEQCKFKSGRECR 422
QY 517 GSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGKYCGVVRKYKSCNLEPCP----- 571
Db 423 ASMSECDPAEHCCTQSS-----ECPADVFNHNGQPC-D---NYGYCYNGNCPIMYHQ 471
QY 572 -----SSASGKSFREOCEAFNGYHNHSTNRLTLAVAWPKYSGVSPRD-KC-KLIC 620
Db 472 CYALFGADVIEAEDSCFKQKNGYGYCRKENGKKTIPCA-----PEDVKCGELYC 522
QY 621 RANGTG-----YFYVLAPK-----VVDGTLCSPDSTSVGVQKCI 655
Db 523 KDNSPGQNNPKMFYSNDEHKGMLPFTKCA--DGKVCVSNHGV 565

RESULT 14

S55270

C:Species: Crotales - western diamondback rattlesnake

C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000

C:Accession: S55270

R:Zhou, J., Smith, J.B., Grossman, M.H.

Biochem. J. 307, 411-417, 1995

A:Title: Molecular cloning and expression of catrocollastatin, a snake-venom protein from

A:Reference number: S55264; MUID:95251603; PMID:773377

A:Accession: S55270

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-609 <2HO>

A:Cross-references: GB:021003; NID:g710353; PIDN:AAC59672.1; PID:g710354

C:Superfamily: mouse meltrin alpha; disintegrin homology

Query Match 6.6%; Score 340.5; DB 2; Length 609;
Best Local Similarity 22.4%; Pred. No. 2.2e-15;
Matches 160; Conservative 100; Mismatches 282; Indels 17; Gaps 32;

QY 1 MLLGLITLAFAGRTA-----GGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDGLIF 55
Db 4 VLLVTICLAAPPYGGSIILSGNVNDYEVYIPKVTALPKG-----AVQPKYEDAMQY 57
QY 56 QITAFQEDFYLHLPDAQFLAPAFSTEHLGVPLQGTGGSSDLRRKCFYSGDVNAEPDSPA 115
Db 58 ELKNGEPVVLHLGKNGKGLFSKDYSETHYSPDGREITTYPLVEDHCYHGRINDADSTA 117
QY 116 AVSLCGGLRGAFGYGAEYVISP--LPNASAPAAQNSQGAHLQRRGVGPGSGDPTSR 173
Db 118 SISACNLKGHFQKQEMYLIEPLKLPDSEAHAVYKENV-----KEDIALKM 166
QY 174 CGVAGWNPAIRALDPYKPRRAGFESRRRSRAKRFVSIIP-----RYVETLVV 225
Db 167 CGVTQNW-----ESYEPIK-----KASQLVWTAEBHQYNPRFVELFLV 205
QY 226 ADESMVKFGADLEHY---LLTLATAARLYRHPSTILNINIVVVKVLLLRDRDSGPKVT 282
Db 206 VDKAWVTKNNGDDLDKTRMYEIVNTVNEYRYMY-----IHVALVGLIEWSNED---KIT 258
QY 283 --GNAALTNRNFCWOKKLNKVDKHPYWDTAIFLTRODLGATTCDTLGMADVGTMCD 340
Db 259 VKPEAGVTLNAGFENKRTDLLTRKKH-----DNAQLLTAIDL-----DRVIGLAYVGSMDCH 309

QY 341 PKRSCSVIED---DGLPSAFTTAHELGHVFNMPHDN-VKVCBEVEFGKLRANHMMSPTLIQ 396
Db 310 PKRSTGIQDYSEINLVAVIMAHMGHNLGINHDSGYSCGDYACIMRPEISPEFSTF- 368
QY 397 IDRANPWSACSAIITDFLDSGHGDCLLDOP-----SKPISLPEDLPGASYTLSSQOCEL 450
Db 369 -----FNSCYFECWDFIMNHNPECILNEPLGTDIISPPVCGNELL-----EVEECD- 416
QY 451 AFGVGSKPCPYQWYCTKLWCTCKAKQWVCQTRHPFPWADGTSCGGKLCCLKGACVERHNLA 510
Db 417 -----CGTPENCQNECCDA-----ATCKLK-----SGSQCGHGD-----CCEQCKF 452
QY 511 NKHRVD--GSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGKYCGVVRKYKSCNLE 568
Db 453 SKSGTECRASMECDPAEHCCTQSS-----ECPADVFNHNGQPCLD---NYGYCYNG 501
QY 569 PCP-----SSASGKSFREOCEAFNGYHNHSTNRLTLAVAWPKYSGVSPRD 614
Db 502 NCPIMYHQCVDLFGADVIEAEDSCFERNQKNGYGYCRKENGKKTIPCA-----PED 552
QY 615 -KC-KLICRANGTG-----YFYVLAPK-----VVDGTLCSPDSTSVGVQKCI 655
Db 553 VKCGRLYCKDNSPGQNNPKMFYSNDEHKGMLPFTKCA--DGKVCVSNHGV 603

RESULT 15

S60257

C:Species: Mus musculus (house mouse)

C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000

C:Accession: S60257

R:Yagami-Hiromasa, T.; Sato, T.; Kurisaki, T.; Kamijo, K.; Nabeshima, Y.I.; Fujisawa-

Nature 377, 652-656, 1995

A:Title: A metalloprotease-disintegrin participating in myoblast fusion.

A:Reference number: S60257; MUID:96026308; PMID:7566181

A:Accession: S60257

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-903 <YAG>

A:Cross-references: EMBL:D50411; NID:gl054586; PIDN:BAA08912.1; PID:gl054587

C:Superfamily: mouse meltrin alpha; disintegrin homology

F.421-503/Domain: disintegrin homology <DIS>

F.349/Active site: Glu #status predicted

Query Match 6.3%; Score 324; DB 2; Length 903;
Best Local Similarity 21.5%; Pred. No. 4.6e-14;
Matches 159; Conservative 73; Mismatches 246; Indels 262; Gaps 27;

QY 2 LLLGLITLAFAGRTA-----GGSEPEREVVPIRLDPDINGRRYYWRGPEDSGQCG 52
Db 15 LLLLAGALLAPRAAGMSLWDQRAYEVARSL--LSKDFCIPGQSI-----PAKDHPDV 68
QY 53 LIFOITAFQEDFYLHLPDAQFLAPAFSTEHLGVPLQ-----GLTGGSSDLRRKCFYSGEV 107
Db 69 LTVQLQLESRLDILSLERNEGLIANGFTETHY---LDQGTVDVSLTRNHTD--HCYHGHV 123
QY 108 NAEPSDPAVSLCGGLRGAFGYGAEYVISP-LPNAS-----APA-AQNSOGAHLQRRG 161
Db 124 OGDAAVSVLSLSCDRLGLIMFENKTYSLPEMKNNTDSYKLVPAESMTNIOGL----- 176
QY 162 VPGPSGDPTSRGCVASGNWPAIRALDPYKPRRAGFESRRRSRAKRFVSIIPRYVE 221
Db 177 -----CGSQHNKSNLTMEDVSP-----GTSQMRARRHK---REFLWKTKVE 215
QY 222 TLVVADESMVKFGADLEHYLLTLATAARLYRHPSTILNINIVVVKVLLLRDRDSGPKV 281
Db 216 LVIVADNREFQKQKDLKLVKQRLIEIANHVDKFPRLN-IRIVLVGVVWVNDIHK-CST 273
QY 282 TGNRAALTNRNFCWOKKLNKVDKHPYWDTAIFLTRODLGATTCDTLGMADVGTACDP 341
Db 274 SODPTRLHEFLDWRKIKLLPRKSH-----DNAQLISGYVFOGT---TIGMAPINSMCTA 325
QY 342 KRSCSVI---EDDGLPSAFTTAHELGHVFNMPHDN-VKVCBEVEFGKLRANHMMSPTLIQ 397

```
Db 326 EQSGVWMDHSDSPLGAATVLAHELGHNFEGMNHDTLFGCSORMAAEKGGCMMNPS----- 381
QY 398 DRANP-----WSACSAIITDFLDSHGDCLLDOP----- 427
Db 382 -TGFPFPMVFSSCRKDLFASLEKGMGMLFNLPEVKQAFGRKCGNGYVEEGEECDGCE 440
QY 428 -----SKPISLPEDLPGAS--- 441
Db 441 PEECTNRCCNATCTLKPDAVCAHGQCCEDCOLKPPGTACRGSSNSCDLPEFCTGTAPHC 500
QY 442 -----YTLQQCELAFGVGSKPCPYMOY----- 464
Db 501 PANVYLHDGHPGQGVGYNGICOTHEQQCVTLWPGAKPAPGICFERVNSAGDPYGNC 560
QY 465 -----CTKLWCTGKAK-----COMVCQTRHEPW 487
Db 561 GKDSKSAFAKELRDAKCGKIQCGGASRPVIGTNAVSIETNIPQOEGGRILCRGTHVYL 620
QY 488 AD-----GTSCGEGKCLKGACVERHNLNKHRYDGSWAKWDPYPCGSRTC-GGGV 536
Db 621 GDDMPDPGLVLAGTKCAEGKICLNRRQNISVFGVHK-----CAMQCHGRGV 667
QY 537 QLARROCTNPTPANGKYCE 556
Db 668 CNNRKNC-----HCE 677
```

Search completed: May 9, 2003, 15:24:10
Job time : 38 secs

GenCore version 5.1.1-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:19:01 ; Search time 45 Seconds
(without alignments)
4349.884 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLGLTLAFAGRTAGGSE.....DOCNLRHPQBLDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP_TREMBL_21.*
- 2: SP_Archaea.*
- 3: SP_Bacteria.*
- 4: SP_Fungi.*
- 5: SP_Invertebrate.*
- 6: SP_Mammal.*
- 7: SP_Mhc.*
- 8: SP_Organelle.*
- 9: SP_Phage.*
- 10: SP_Plant.*
- 11: SP_Rodent.*
- 12: SP_Virus.*
- 13: SP_Vertebrate.*
- 14: SP_Unclassified.*
- 15: SP_Rvirus.*
- 16: SP_Bacteriap.*
- 17: SP_Archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5162	100.0	950	Q8TE58	Q8TE58 homo sapien
2	1695	32.6	340	Q91256	Q91256 mus musculus
3	1587	30.7	2165	Q19791	Q19791 caenorhabdi
4	1295.5	25.1	1081	Q8TE60	Q8TE60 homo sapien
5	1247.5	24.2	1072	Q8TE57	Q8TE57 homo sapien
6	1239	24.0	1688	Q8SXB0	Q8SXB0 drosophila
7	1220	23.6	1095	Q8TE56	Q8TE56 homo sapien
8	1206.5	23.4	1229	Q9VF61	Q9VF61 drosophila
9	1185.5	23.0	1054	Q9W493	Q9W493 drosophila
10	1165.5	22.6	1207	Q8TE59	Q8TE59 homo sapien
11	1122	21.7	1223	Q8WXS8	Q8WXS8 homo sapien
12	1120.5	21.7	1159	Q8TEY8	Q8TEY8 homo sapien
13	1118	21.7	1223	Q8TE55	Q8TE55 homo sapien
14	893.5	17.3	1427	Q96L37	Q96L37 homo sapien
15	830	16.1	269	Q9GL54	Q9GL54 oryctolagus
16	706	13.7	1280	Q9EPX2	Q9EPX2 mus musculus

17	693	13.4	1235	4	Q95428	Q95428 homo sapien
18	681	13.2	2167	5	Q76840	Q76840 cu-norhabdi
19	663.5	12.9	1572	5	Q44938	Q44938 laemochus
20	599	11.6	1062	5	Q19204	Q19204 caenorhabdi
21	581.5	11.3	2174	5	Q9GQRO	Q9GQRO drosophila
22	579.5	11.2	3060	5	Q9AV4	Q9AV4 drosophila
23	569.5	11.0	766	4	P82987	P82987 homo sapien
24	559.5	10.8	1444	5	Q17591	Q17591 caenorhabdi
25	558	10.8	192	6	Q5N24	Q5N24 equus caball
26	557	10.7	3198	5	Q9U8G8	Q9U8G8 manduca sex
27	550.5	10.7	525	4	Q96RW4	Q96RW4 homo sapien
28	518	10.0	187	6	Q95N23	Q95N23 equus caball
29	511.5	9.9	1091	5	Q9W126	Q9W126 drosophila
30	510.5	9.9	872	5	Q22580	Q22580 caenorhabdi
31	493	9.6	951	4	Q60345	Q60345 homo sapien
32	487.5	9.4	1014	5	Q95R33	Q95R33 drosophila
33	426	8.3	790	5	Q8T458	Q8T458 drosophila
34	411.5	8.0	1059	5	P90884	P90884 caenorhabdi
35	410	7.9	183	11	Q99JP1	Q99JP1 mus musculus
36	402.5	7.8	364	4	Q9UG01	Q9UG01 homo sapien
37	379	7.3	899	13	Q8UVF1	Q8UVF1 coturnix co
38	365	7.1	922	13	Q8UVF2	Q8UVF2 coturnix co
39	356	6.9	610	13	Q93523	Q93523 bothrops ja
40	356	6.9	610	13	Q8QG88	Q8QG88 bothrops in
41	343	6.6	610	13	Q9YI20	Q9YI20 agkistrodon
42	343	6.6	622	4	Q9H8X0	Q9H8X0 homo sapien
43	340.5	6.6	609	13	Q90282	Q90282 croctalus at
44	340.5	6.6	609	13	Q9W6M5	Q9W6M5 agkistrodon
45	338	6.5	192	11	Q9CX59	Q9CX59 mus musculus

ALIGNMENTS

RESULT 1

Q8TE58 PRELIMINARY; PRT; 950 AA.

AC Q8TE58;

DT 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Metalloprotease disintegrin 15 with thrombospondin domains.

GN ADAMTS15.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21856482; PubMed=11867212;

RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,

RA Lopez-Otin C.;

RT "Cloning, expression analysis, and structural characterization of

seven novel human ADAMTS, a family of metalloproteinases with

disintegrin and thrombospondin-1 domains.";

RL Gene 283.49-62(2002).

DR EMBL; AJ315733; CAC86014.1; .

KW Integrin; Protease.

SQ SEQUENCE 950 AA; 103286 MW; 5DFBE18285CCCC3B CRC64;

Query Match 100.0%; Score 5162; DB 4; Length 950;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLTLAFAGTAGSGPEREVVPIRLDPDINGRRYRWGPDGSGQGLIFQ11AF 60

Db 1 MLLGLTLAFAGTAGSGPEREVVPIRLDPDINGRRYRWGPDGSGQGLIFQ11AF 60

QY 61 QEDFYHLTPDAQFLAPAFSTEHLGVLPGLTGSSDLRRCFYSGDVNAEPDFAAVSLC 120

Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVLPGLTGSSDLRRCFYSGDVNAEPDFAAVSLC 120

QY 121 GGLRGAFYRGAEYVISPPLNAPAAQRNSQGAHLQRRGVPGPGSDPTSRGCVASGW 180

Db 121 GGLRGAFCYCAEYVLSPLPNASAPQAQRNSQGAHLQLRGVPGSPGDPSTSCGVASGW 180
Qy 181 NPAILRALDPKPRRAGGSRSSRRGRKRRFVSIPRYVETTLVVADESVMKPHGADLEH 240
Db 181 NPAILRALDPKPRRAGGSRSSRRGRKRRFVSIPRYVETTLVVADESVMKPHGADLEH 240
Qy 241 YLLTLLATAARLYRHPSTLNPINWVVKVLLLRDRDSCPKVTGNAALTLLNFCAWOKKLN 300
Db 241 YLLTLLATAARLYRHPSTLNPINWVVKVLLLRDRDSCPKVTGNAALTLLNFCAWOKKLN 300
Qy 301 KVSQKHPEYWDATILFTRODLGGATTCDTLGMADVGTMCDFKRSVCVIEDGLPSAFTTA 360
Db 301 KVSQKHPEYWDATILFTRODLGGATTCDTLGMADVGTMCDFKRSVCVIEDGLPSAFTTA 360
Qy 361 HELGHVFNMHDNVKVCBEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420
Db 361 HELGHVFNMHDNVKVCBEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420
Qy 421 DCLLDQPSKPTSLPDLPGASYTLISQCELAFAFGVSKPCPYMQYCTKLWCTGKAKGOMVC 480
Db 421 DCLLDQPSKPTSLPDLPGASYTLISQCELAFAFGVSKPCPYMQYCTKLWCTGKAKGOMVC 480
Qy 481 QTRHPFWADGTSQSGKGLKCLGACVERHNLNKHRYVDSWAKWDYGPCSRCTCGGVQOLAR 540
Db 481 QTRHPFWADGTSQSGKGLKCLGACVERHNLNKHRYVDSWAKWDYGPCSRCTCGGVQOLAR 540
Qy 541 ROCTNPTFANGKCYCEGVRYKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTTLA 600
Db 541 ROCTNPTFANGKCYCEGVRYKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTTLA 600
Qy 601 VAWVPKYSVSPDRCKLICKRANGTGYFYVLAPKVVVDGTLCLSPDSTSVCVQKCIKAGCD 660
Db 601 VAWVPKYSVSPDRCKLICKRANGTGYFYVLAPKVVVDGTLCLSPDSTSVCVQKCIKAGCD 660
Qy 661 GNLGSKKRRFDCGVGCGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRYKGLI 720
Db 661 GNLGSKKRRFDCGVGCGDNKCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRYKGLI 720
Qy 721 GDNTYALKNSGKYLNGHFWVSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLT 780
Db 721 GDNTYALKNSGKYLNGHFWVSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLT 780
Qy 781 VEVLVSGKMTPPRVRYSYFLPKPEPREDKSSHPKDPGRGSPVLHNSVLSLSNOVEQDDRRPP 840
Db 781 VEVLVSGKMTPPRVRYSYFLPKPEPREDKSSHPKDPGRGSPVLHNSVLSLSNOVEQDDRRPP 840
Qy 841 ARWVAGSWGPCASGSGLOKRAVDCRSAGQRTVPACDAAHRPVTQACGECPTWELS 900
Db 841 ARWVAGSWGPCASGSGLOKRAVDCRSAGQRTVPACDAAHRPVTQACGECPTWELS 900
Qy 901 ANSPCKSCGGRGFQRRSLKCVHGGGRLLARDQCNLHRRKPOELDFCVLRPC 950
Db 901 ANSPCKSCGGRGFQRRSLKCVHGGGRLLARDQCNLHRRKPOELDFCVLRPC 950

RESULT 2
Q91256 PRELIMINARY; PRT: 340 AA.
AC Q91256:
Q91256:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Similar to a disintegrin and metalloproteinase with thrombospondin
DE motifs 1 (ADAMTS-1) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC009667.1; AAH09667.1; -.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 2.
KW Integrin.
FT NON_TER 1
SQ SEQUENCE 340 AA; 37021 MW; 06562C747634A8BD CRC64;
Query Match 32.6%; Score 1685; DB 11; Length 340;
Best Local Similarity 91.7%; Pred. No. 3.4e-133;
Matches 308; Conservative 11; Mismatches 17; Indels 0; Gaps 0;
Qy 615 KKLICRANGTGYFYVLAPKVVVDGTLCLSPDSTSVCVQKCIKAGCDGNLGSKKRFDKCGV 674
Db 5 KKLICRANGTGYFYVLAPKVVVDGTLCLSPDSTSVCVQKCIKAGCDGNLGSKKRFDKCGV 64
Qy 675 CGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQYKGLIGDNYLALNSOGK 734
Db 65 CGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQYKGLIGDNYLALNSOGK 124
Qy 735 YLLNGHFVYSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLTVEVLVSGKMTPPRV 794
Db 125 YLLNGHFVYSAVERDLVVKGLSLLRYSGTGTAVESLQASRPILPLTVEVLVSGKMTPPRV 184
Qy 795 RYSFYLPKPEPREDKSSHPKDPGRGSPVLHNSVLSLSNOVEQDDRRPPARWVAGSWGPCSAS 854
Db 185 RYSFYLPKPEPREDKSSHPKDPGRGSPVLHNSVLSLSNOVEQDDRRPPARWVAGSWGPCSAS 244
Qy 855 CGSGLQKRAVDCRSAGQRTVPACDAAHRPVTQACGECPTWELSAWSPCKSCGGRGFQ 914
Db 245 CGSGLQKRAVDCRSAGQRTVPACDAAHRPVTQACGECPTWELSAWSPCKSCGGRGFQ 304
Qy 915 RRSCLKCVHGGGRLLARDQCNLHRRKPOELDFCVLRPC 950
Db 305 RRSCLKCVHGGGRLLARDQCNLHRRKPOELDFCVLRPC 340

RESULT 3
Q19791 PRELIMINARY; PRT: 2165 AA.
AC Q19791; Q27524;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadaty S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lighning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smailson J., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans";
RL Nature 368:32-38(1994).
DR EMBL; Z69361; CAA93288.1;
DR EMBL; Z69360; CAA93288.1; JOINED.
DR EMBL; Z69360; CAA93287.1;
DR EMBL; Z69361; CAA93287.1; JOINED.

[illegible]

```
QY 311 DTAIFTRQDLGCG--ATTCDTLGMADYGTMCDDPKRSCSVIEDDGLPSAFTTAHELGHVFN 368
DB 333 DHAILLTGFDLCSWKNKPCDPLGAPTSGMCKYRSCTINEDTGLGLAFTIAHESGHNGF 442
QY 369 MPHDNV-KVCEVFGKLRANHMSPTLIQIDRANPWSACSAAITDPLDSGHGDCLLDQP 427
DB 443 MHDGEGNPKRKAEG-----NIMSPITLTGNGVFSWSSCSQYLLKFLSTPOAGCLVDEP 497
QY 428 SK--PISLPEDLPASYSYLSQOCELATGVGSKPCP---YMOYCTKLWC--TGRAKGQMYC 480
DB 498 KOAGQYKPDPLPGQIYDADTQCKWQFGAKAKLCSLGFVKDICKSLMCHRVGHR-----C 552
QY 481 QTRHPWADGTSCGEGKLCACACVERHNLNKHVRVDGSAWAKWDPYGPCSRTCTGGGVOLAR 540
DB 553 ETKFMPAAGTGVGLSMWCROGQCVKTELGPRPIHGWSAWSKSECSRTCTGGGVGFQ 612
QY 541 ROCTNPTPANGKGYCEGVVRYRSCNLEPCSSASGSKSFREEQCEAFN-----GYNISTN 595
DB 613 RHNCNPKQYGGIFCGSSRIYQLCINPNC--NENSLDFRAQQAQAEYNSKPFGRWFYQ-- 668
QY 596 RLTLAVANVPKYSGVSPRDKCLICRANGTGYFYVLAPKVVVDGTLCSPDSTSVCVQKCI 655
DB 669 -----WKP-YTKVEEDRCKLYCKAENFEFFAMSGKVKGDTGPCSPNKNDVCIDGVC 720
QY 656 KAGCDGNLGSKKRFDKCGVCGDNKSKKVTGLFTK--PMHGYNFVVAIPAGASSIDIRQ 713
DB 721 LVGCDHELGSRAVSDACGVCKGDNSTCKFYKGLYNQHKANEYYPVVLIPAGARSIEIQ 780
QY 714 RYKGLIGDDNYLAKNSQGYLLNGHVVSVERDLVVKGSLRLYSGTGTAVESLOASR 773
DB 781 -----LOVSSSYLLAVRSLSQYLYTGGWSID-WPGEFFPAGTTFEYORSFNRPERLYAPG 834
QY 774 PILEPLTVEVLSVGKMTPPRVRYSPYLPK-----EPREDK-----SSHPKDPGRPSV 820
DB 835 PNEITLVFEILLMOGK--NPGTAWKALPKVNGTPPATKPAYTWSIVQSECVSCGGY 892
QY 821 LHNVSLSLSNOVEQ-----PDRP-----PARWVAGSGPCASCSGSLQK 861
DB 893 INVKAICLRDQNTQVNSSFCSAKTPVTEPKICNAFSPAYWMPGEWSTCSKSCAGGQOS 952
QY 862 RAVDCRGSAGORTVA-----CDAHRVETQAC-GEPC-PTWELSAWSPKSCGGRGFO 914
DB 953 RKIQCVQKPKPQKEBAVLSLCPVS--TPTQVQACNSHACPPQWSLGPWSQCSKTCGRGVR 1011
QY 915 RRLSKCVHGGRLLRDQCNLHRKPQELDFCVLRPC 950
DB 1012 KRELCKGSAETLPSQCTSLPRELQEGCVLGR 1047

RESULT 5
Q8TE57 PRELIMINARY; PRT; 1072 AA.
AC Q8TE57
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Metalloprotease disintegrin 16 with thrombospondin type I motif.
GN ADAMTS16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002).
DR EMBL; AJ315734; CAC86015.1; -;
```

```
KW Integrin; Protease.
SQ SEQUENCE 1072 AA; 119491 MW; 081EEFF78F47D061 CRC64;
Query Match 24.2%; Score 1247.5; DB 4; Length 1072;
Best Local Similarity 32.9%; Pred. No. 1.2e-95;
Matches 332; Conservative 126; Mismatches 386; Indels 165; Gaps 38;
QY 51 QGLIFOITAFQEDFYHLTPDQAFAPAFSTFHLGVPLQGLTGGSS-----DLRRCFY 103
DB 95 ESLHLRLKGRDFHMDLTSSSLVAPGFIQTL-----GKTGTSVQTLPPEDF--CFY 147
QY 104 SGDVAEPDPSFAVSLCGLRGAFGYRGAEYVISPLPNASAPAAOANSOG-----AHLQRR 160
DB 148 QGSLSRHNSVALSTFCQGLSGMIRTEADYFLRPLPSHLKGLRAAQSPSHVLR 207
QY 161 G-----VPGGPGSDPTSR-----CGVASGWNPA-----LIRA 187
DB 208 STEPHAPGASEVLVTSRTWELAHQPLHSSDLRLGLPQKHFCGRRKKYMPPOPPKEDLFL 267
QY 188 LDPYKPRRAGFESRRRRSGRAKRVSTPRVETLVVADESMVKFPG--ADLEHYLLILL 246
DB 268 PDEYK---SCLRHKRLSLRSHRNEEL-----NVETLVVYDKMMQNHGENTITVLTIL 319
QY 247 ATAARLYRHPSTILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAMQKKL-NKVS DK 305
DB 320 NMVSALFKDGTIGGNINIAIVGLILLEDPQGLVISHHADHTLSFCQWQSGLMGKDGTR 379
QY 306 HPEYWDTAILFTQRDLGCG--ATTCDTLGMADYGTMCDDPKRSCSVIEDDGLPSAFTTAHEL 363
DB 380 H-----DHAILLTGLDICSWMKNEPCDPLGAPISGMSKYRSCTINEDTGLGLAFTIAHES 435
QY 364 GHVFNPHDNV-KVCEVFGKLRANHMSPTLIQIDRANPWSACSAAITDPLDSGHGDC 422
DB 436 GHNFIMHDEGHNMKKSEG-----NIMSPITLAGRNGVFSWSPCSQYLUKFLSTAQALC 490
QY 423 LLDQPSKPI---SLPBDLPASYSYLSQOCELATGVGSKPCPY---MQYCTKLWC--TGXA 474
DB 491 LADQP-KPYKEYKYPEKLPGLDYDANTQCKWQFGEKAKCLMDLDFKDKALWCHRIGRK 549
QY 475 KGQMVQOTRHPFWADGTSCEGKLCUKGACVERHNLNKHVRVDGSAWAKWDPYGPCSRTCGG 534
DB 550 -----CETKEMPAAEGTICGHDWCRGGQCVKYGDEGPKPTHGHWSDWSWSPCSRTCGG 604
QY 535 GVQLARQCTNPTPANGKGYCEGVVRYRSCNLEPCSSASGSKSFREEQ-----EAFNG 589
DB 605 GVSHRSRLCTNPKPSHGKFCFCEGSTRTLKLCNSQKCPDS--VDFRAAOCAEHSRRFSG 662
QY 590 YNHSTNRLTLAVANVPKYSGVSPRDKCLICRANGTGYFYVLAPKVVVDGTLCSPDSTSV 649
DB 663 RHYK-----WKP-YTQVEDQDLCKLYCIAEGDFEFSLSNKYKDGTPCSEDSRNVC 712
QY 650 VQKCIKAGCDGNLGSKKRFDKCGVCGDNKSKKVTGLFTKPMH--GYNFVVAIPAGAS 707
DB 713 IDGICERVGCDNLGSDADEVGCGVCGNNSACTIHRGLYTKHHHTNQYHYHMVTIPSGAR 772
QY 708 SIDIRQRYKGLICDDNYLAKNSQGYLLNGHVVSVERDLVVKGSLRLYSGTGTA 767
DB 773 SIRIYENNV-----STYSIISRNALRYLLNGHWTDWQGR-YKFGTTFDYRRSINEPE 826
QY 768 SLOASRPILPLATVEVLSVGKMTPPRVRYSYFLPKFREDKSKHPKDPGRPS-----819
DB 827 NLATGPTNETLIVELLFQGR--NPGVAWEYSMPR-----LGTEKOPPAQPSYTAIVRS 879
QY 820 -----VLHNSVLSLNOVE-----QPDPRP-----PARWVAGSWGPC 851
DB 880 ECSVSCGGQMTVREGCYRDLKFQVNMNSFCNPKTRPVTLGVPCVKSACPPSWSVGNMSAC 939
QY 852 SASCGSLQKRAVDC--RGSAGORTVPA--CDAHRVETQAC-GEPC-PTWELSAWSPC 905
DB 940 SRTCGGAQSRPVOCTRRVHYDSEVPASLCPQP-APSSRQAACNSQSCPPAWAGWAEC 998
QY 906 SKSGRGFRRLSKCVGHG-----GRLIARDQCNLHRKPQELDFCVLRPC 950
```

Db 999 SHTCGKWRKRAVACKSTNPSARAQLLPDAVCTSEPKRMHEACLQRC 1047

RESULT 6

Q8SXB0 PRELIMINARY; PRT; 1688 AA.

AC Q8SXB0; DB 5; Score 1239; DB 5; Length 1688;

DT 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DE GH16393p.

CG6107.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidae; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nuncio J., Pacieb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Celniker S.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY094716; AAM1069.1;

SQ SEQUENCE 1688 AA; 189867 MW; 48FB8DD4DE0CA4D2 CRC64;

Query Match 24.0%; Score 1239; DB 5; Length 1688;

Best Local Similarity 32.0%; Pred. No. 1.2e-94;

Matches 321; Conservative 142; Mismatches 339; Indels 200; Gaps 39;

Qy 55 FOITAFQEDFVLTLPDPAFLAPAFSTHGLVPLQG-----LTGSSDLRR--CFYSGD 106

Db 296 YNLNVFGQLHLVLRQDASF--HNHSMTHRIKLEGEHPGTEAEQRIHLGCFVSGY 354

Qy 107 VNAEPDSFAAVSLCGGLRGAFYGAAYVISPFPNASAP-----AAQNSOGAHLQ 158

Db 355 VEDDPHSMVSVSLCGGTGIKTSFGALLIQPNRTSSDEVLRHVPKRSQNRN--HAYS 412

Qy 159 RRGVPGSGDPTSRGCVASGNWPAILRALDPYKPRAGFGE-----SRSRRSRGAKRFV 214

Db 413 KFEL-----GLDDFMKSLQVQEEQKSKRKNRKHVA 448

Qy 215 SIPRYV---ETLVVADESVMKFFGADLEHYLLTLTAAFLRHPSTLNPINIVVVKVLL 271

Db 449 DVDNQVTVLEVLAVDSMKQFHGEDLPYLLILMSIVSFADASIGNSIRILLVRLIS 508

Qy 272 LRDRDSGPV---TGNAAFLRNFCAWQKKLVSKHPYWDATLFTQDLGCGAT--- 325

Db 509 L-----PNINDOTHSSNEMLKHFQCP---INQSGYER---DTAMLTITREPICGSPVGK 555

Qy 326 TCDPLGMADVGTCDPKRSCSVIEDGLPSAFTTAHELGHVFMHNDNVKVCVEVFGKLR 385

Db 556 ICHMLGAEILGTVCSSSSCSIVODTGLPTAFTMAHELGHILNMHDDDDKCMPPYVTRON 614

Qy 386 AN---HMSPTLIQIDRANPWSACSAAITDLDSDHGDCLLDQPSKPSLSP---EDLPG 439

Db 615 NNKVLHIMSSVMGTHMHPWSKSRHFVSEFLEKTDKSL---ETSVGAHPIYGTRELPG 672

Qy 440 ASYLSQOCEIARGVGSKPCPYQYCTKLWC-----TGKAKQMVCOYRHFHWADGTSCG-E 495

Db 673 EYISLDAQOCLSFQNDGFCYPTDECKRLWCNRTSGNSNEQ---CASSNLPWADGTGCGSS 730

Qy 496 GKLLKGACVERHNANKH-----RVDGSWAKWDPIGFCSCRTCCGGVGLARQCTNPTPANG 551

Db 731 GHWCQKCKVS---NKHGTVGVNGWGWPTFTFCSLTCCGGVGSRRCCNQPVPENG 786

Qy 552 GKCEGVVRYKSCNLEPCPSSASGSKSFRFEOCEAFNGYNHSTNRLTLAVAWPKYSGVS 611

Db 787 GKYGTSRKKYRSCNTHQCPGSGMDP---REQQCYANGRNMNIPGVNPDTKWVPKYE--- 841

Qy 612 PRDKCKLICRANGTGYFYVLAPKVVDTGLCSPDSTSVCVQKCIKAGCDNGLSKKRFDK 671

Db 842 -KDACKLFCRMDMKVTFYMLKSMVTDGTSCAVDSFDKCVNGICRPAGCDNELNSIAKLDK 900

Qy 672 CGVCGDNKSKKVTG-LFTKPMHGYN-----FVVAIPAGASSIDIRGKGLI 720

Db 901 CGVCEGRNDTCHEVTGNLLVNSLLGLNDGNEPNKTIYYVTRIPKGSNIITQGYYP--- 957

Qy 721 GDDNYLALKNSQKYLNLGHFVVSVERDLVVGSLRLYSVGTGTAVESLQA---SRPILEP 778

Db 958 -DQNFIVLTDNRDNLNGKF-LKTYPLKFVYAGVTQMOTGSSSVVEQVNTIYSWKLSD 1015

Qy 779 LTVFVLSVGMKTPPR-----VRSFYLPKPRE- 806

Db 1016 LIVQIIISLDVSPSKRQDTVLLSYTYIDKPPDYAEVEIYRWEMQAPSNCDSLCGRSHR 1075

Qy 807 -----DKSSHK-DRGPSV---LHNSVLSLSNQVQPDPPARVW 844

Db 1076 LPACISTTQGVKVAPOFCDKSAKPKIDDRACNTDCLRLNLTVTISIE----- 1121

Qy 845 AGSWGPCASCCGS-GLQKRAVDC--RGSAGQRT---VPACDAHRPVTQACGEPCTW 897

Db 1122 -----CSAACGELGTREKTYACVQFTTNMORSNIVDMYSYCKLFDVAYHEECREGC-W 1173

Qy 898 ELSAWSPCKSCGRCGFORSLKCVGHGGLLARDQCNLHRKP 939

Db 1174 VLSEWSTCKSCGTSQQAQREAHCHLNSR-VSDDLGNPRTKP 1214

RESULT 7

Q8TE56

ID Q8TE56 PRELIMINARY; PRT; 1095 AA.

AC Q8TE56;

DT 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Metalloprotease disintegrin 17, with thrombospondin domains.

GN ADAMTS17.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21856482; PubMed=11867212;

RA Cal S., Odaya A., Llamazares M., Garabaya C., Quesada V.,

RA Lopez-Otin C.;

RT "Cloning, expression analysis, and structural characterization of

RT seven novel human ADAMTSs, a family of metalloproteinases with

RT disintegrin and thrombospondin-1 domains.";

RL Gene 283:49-62(2002).

DR EMBL; AJ315735; CAC86016.1; -.

KW Integrin; Protease.

SQ SEQUENCE 1095 AA; 121099 MW; A5C6540484754D5D CRC64;

Query Match 23.6%; Score 1220; DB 4; Length 1095;

Best Local Similarity 31.6%; Pred. No. 2.5e-93;

Matches 331; Conservative 139; Mismatches 386; Indels 190; Gaps 40;

Qy 2 LLLGLTILAF-----AGRTAGGSEPEREVVPIRLDPDI-----NGRRVYWRGPE 46

Db 10 LVLPVLLLVWGLDPTAVGDAADVEVLPWRVPPDVLPLPAAPCPRRRRRPTTP 69

Qy 47 DS-----GQGLIFQITAFQEDFYHLTPDQAFAPAFSTHGLVPLGGLTGGSSDLRRC 101

Db 70 AAPRAPGERALLHLPAFGRLDYQLRLRDLRFLSRGFEVEEAGAARR--RGRPAEL--C 125

Qy 102 FYSGVDPNPEPSFAVSLCGGLRGAFG--YRGAEVV-ISPPLNASAPAAQNSOAHLLQ 158

Db 126 FYSGRVLHGPGSLVSLACGAAGGLVGLIQGLQEQEVLQPLNNSGQFSGRE---HLIR 181

Qy 159 RR-GVPGGPGSD---PTSRCGVASGNWPAILRALDPYKPRAGFGE-SRSRRSRGAKRF 213

Db 182 RWSLTSPSAEAQRPEQLCKVL-----EKKPTWGRSDWRERRAIRL 228
Qy 214 VSPRYVETLVVAVESWVKFGAD-LEHYLLTLATARLYRHPISILNPINIVVKKVLL 272
Db 229 TS-EHTVETLVVADAMVQYHGAEAARFILTVMNVMYMFQHOSLGKINQIVTKVLL 287
Qy 273 RDRSGPKVTGNAALTANFCAWKK-----LNKV--SDKHPEYWDTAIFLTRQDL 321
Db 288 RORPAKLSIGHGERSLESFCHWNEEYGGARYLGNQVPGCKDDPLVDAAEVETRTDF 347
Qy 322 C-GATTCDTLGMADVGMCDPSCSVCIEDGLPSAFTTAHELGHVFNPMHNKVCCEE 379
Db 348 CVHKDEPCDTGVIAVLGGVCSAKRCKVLAEDNGLNLAFTIAHELGNLGMNHD-----DD 402
Qy 380 VFGKLRANMSPTLIQIDRANP-----WSACSAALITDLDGSHGDCLL-----DQSPKIS 432
Db 403 HSSACGRSHIMSGEYV--GRNPDLSSWSSCRDLENFLKSKYSTCLLVTDPRSQTVR 460
Qy 433 LPEDLPGASYTLQOCELAFGVGSKPCPYMAY--CTKLWCTGKAGQWCVCTRFPWADG 490
Db 461 LPHKLPGMHYSANEOQILFGMNATFCRMEHLMCAGLWCL--VEGDFSCCKLDPPLDG 518
Qy 491 TSCGEGKLCGLGACVERNLNKHVRVDGSKWAKWDVPGPCSRCTGGGVQLARQCTNPTPAN 550
Db 519 TECGADKWCRCAGECVSKTPIEH-VGDWSPWGAWSMCSRTCGTGAREPRQRKCDNPPPGP 577
Qy 551 GKCYGCVRVKYSNLCPSASSKSFREOCEAFNGYHSTNRLTLAVAWPKYSGV 610
Db 578 GTHCPGASVEHAGVENLPCPKGL--PSFRDQOQAHDRLPKPKGLTAV-----V 627
Qy 611 SPRDKCKLICRANGTYGYFVLAPKVVDTGLCSPTDSTSVCGVKCICAGCDNGLSKKRF 670
Db 628 VDDKPELYCSPLGKESPLVADVRLDGTGPGPYETDLCVHGCKQICGCDGLIGSAKED 687
Qy 671 KGVCGGDNKSKVYGLFTKPMHGNFVVAIPAGASSIDIRQYKGLIGDNDYLAKN 730
Db 688 RGVCSGDKTKHLVKGDFSH-----ARGTALKDSG-KGSI-----N 723
Qy 731 SOKYLLNGHFVVSVERDLVVKGLRYSNGTAVESLOSARPLEPLTVEVL-----S 785
Db 724 SOWKIPEGEFOIA-----GTVRVVRKG-LWEKISAKGPKLPELHMLVLLFHDQD 773
Qy 786 VGMTPPRVRYSYFLPKPEKDRKSHKDPDR-----GPSV-----820
Db 774 YG-----IHYEYVTVNRTAENQSEPKQDLSLFTWTHSGWEGCSVQCGGERRTIVSC 827
Qy 821 --LHNSVLSLNOVEQDPRPP-----ARWAGSWGSCSASCGLQKRAVD 865
Db 828 TRIVNKTTLVNDSDCPOASPEPEQVRRGNLHPCQSRWVAGPWSPCSATCEKGFQHREVT 887
Qy 866 C----RGSAGORTVPACDAARHPVETQAC-GEPC-PTWELSAWSPCSKSGRGFORSLIK 919
Db 888 CVYQLQNGTHVATRPILYCGPRPAVQSCGQDCLSLNEASEWSQCSASCCKGVWKRIVA 947
Qy 920 CVHGGRLIARQCNLHRKPQELDFC 945
Db 948 CTNSQCK-----CDASTRPRAEEAC 967

RESULT 8
Q9VF61

ID Q9VF61 PRELIMINARY; PRT: 1229 AA.
AC Q9VF61;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE C66107 protein.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.E.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer J.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Betens P.V., Berman B.P., Bhandari D., Brottier P.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies F.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yac Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003709; AAF55199.1; .
DR FlyBase; FBgn0038340; CG6107.
DR InterPro; IPR002870; Pep_M12B_propel.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPl.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01562; Pep_M12B_propel; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 2.
DR SMART; SM00209; TSPl; 3.
DR PROSITE; PS00215; ADAM_MEPRO; 2.
DR PROSITE; PS00092; TSPl; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc.
SQ SEQUENCE 1229 AA; 138136 MW; C1293180AB19B700 CRC64;

Query Match 23.4%; Score 1206.5; DB 5; Length 1229;
Best Local Similarity 32.0%; Pred. No. 4e-92;
Matches 321; Conservative 140; Mismatches 329; Indels 213; Gaps 43;

Qy 55 FOITAFQEDFYHLTPDAQFLAPAFSTHEHGVPLQG-----LTGSSDLRR--CFYSGD 106
Db 195 YNLNVFGRQLHLVLRQDASFV-HNHSMTHIRILKEGEEHPGTEAEARHLCGFYSGY 253
Qy 107 VNAEPDGFAAVSLCGGLRGAFYGAETVISPFPNASAP-----AAQNSQGAHLQ 158
Db 254 VEDDPHSMVSVSLCGGMTGYIKTSFGALLIQPVNRTSDEVLHVRFKSQNRAR--HAVS 311
Qy 159 RRGVPGSGDPTSRGCVASGNPAILRALDPYKPRAGFGE-----SRSRRRSGRAKRFV 214

Db 749 GYGLDICANGRCQKVGCDGLGLSLAREDHGVCNGKCKIIKGFNHTRGAGYVEVL 808
QY 701 AIPAGASSIDIRORGYKGLIGDNDYALKNSOGKYLINGIFVSAVEREDLVKSLRLYS 760
Db 809 VIPAGARIKVVEEK-----PAHSYLAURDA-GKOSINSWKI-----EHS 848
QY 761 G-----TGTA-----ESLAQRPILEPLTVEYL-----SVGKMTPPRVRYFYLPKE 803
Db 849 GAFNLAGTTHVYRRGLWEKISAKGPTTAPLHLVLLFQDQNYG-----LHYEYTPSD 902
QY 804 P-REDKSHKDP-----RGPVHLNSVLSL----- 829
Db 903 PLPENOSSKAPELPMWTHTSWEDCDATCGGGERKTTSCTKINSKISVDNECKKYL 962
QY 830 ---NOVEQPDPRP-PARVWAGSWGPCASCGSLQKRAVDGRSGAGORTVPACD-----AA 881
Db 963 KPEQIRKCNQPCOTRWMTETWPCSTCKGQSGROVACTQQLSNGTLIRARERCIG 1022
QY 882 HRPVETQAC-GEPCPT-WELSAWSPCKSCSGRGFRKSLKCVGHGKLLARDQCNLHRKP 939
Db 1023 PKPASAORCEQDCMTVNEAGWSEFSVKCGKGIHRTVRCNT-----PRKCVLSTRP 1076
QY 940 QELDFC 945
Db 1077 REAEDC 1082
RESULT 11
Q8WXS8 PRELIMINARY; PRT; 1223 AA.
ID Q8WXS8
AC Q8WXS8
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE A disintegrin-like and metalloprotease with thrombospondin type 1
DE motif 14 precursor.
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638061; PubMed=11779638;
RA Bolz H., Ramirez A., von Brederlow B., Kubisch C.;
RT "Characterization of ADAMTS14, a novel member of the ADAMTS
RT metalloproteinase family";
RL Biochim. Biophys. Acta 1522:221-225(2001).
RE EMBL; AF386666; AAL40229.1; -;
DR InterPro; IPR002870; Rep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Rep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp_1; 4.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS02015; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 1.
KW Signal; Integrin; Protease; Metalloprotease.
FT SIGNAL 1 22 POTENTIAL.
SQ SEQUENCE 1223 AA; 133871 MW; 36394AC4D92F17DF CRC64;
Query Match 21.7%; Score 1122; DB 4; Length 1223;
Best Local Similarity 29.9%; Pred. No. 5e-85;
Matches 306; Conservative 122; Mismatches 379; Indels 216; Gaps 40;
QY 45 PEDSDQGLIFQITAFQEDFYHLTPDAQFLAPAFSTEH-----LGVPLQGLTGSSD 97
Db 98 PGRVGRHSLYENVTVFGKELHLRLRPNRLVVPVPGSSVEMQEDREFLRQPLR----- 149
QY 98 LRRCFYSDVNAEPDPSFAANVSLGGLGAFGYRGAETVISPPLNAPAPAAQNSQ-AHL 156
Db 150 -QECVYTGVTGMPGAAVAINSCDGLAGLIRTDSTDFIEPLERQ-----QKEASGRTHV 205

QY 157 LQRRGVP-----GPGSDPTSRGCVASGNPAILRALDPYKPRAGFGESESRSSRSKAKR 212
Db 206 YVREAVQOEWAEPDGDHLNE-AFGLDLPNLLGLVGDO-----LGTERRRHKAPGS 258
QY 213 FVSIPRVETLVVADESMVKFHGAD-LEHYLLTLTAARLYRHPHSILNPINIVVVKVLL 271
Db 259 Y-----SIEVLLVDDSVVRFHGKEHVQVVLTMIVDEIYHDESGLVGHINIALVRLIM 313
QY 272 LRDRDSGPV-TGNAALTIRNFCAWOKKLNKYSKHIPEYWDTAITLFTRODLCGATTCDTL 330
Db 314 VGYRSLSLIERGNPSRSLQEVCRWAHSOORQDPSHAHHHDHVVFLTRQDF-----GPS 367
QY 331 GMADVGTCDPKKCSVIEDDGLPSAFTTAHELGHGVFNPHD---NVKCEEVFGKIRAN 387
Db 368 GYAPVTGMCHPLKSCALNHEDGFSSAFVTAHETGHMEHGGQGCDAETSLGS----- 423
QY 388 HNMSTLIQIDRANPNWASACSAIITDLSHGDCILDDQSPKI-SLPEDLPCASVTLQ 446
Db 424 -VMAPLVQAAFRHFWRSCKLSRYLPS--YDCLDDPDPAPWPQPELPGINYSMDE 480
QY 447 QCELATGVGSKPC---PYMYYCTKLWCTGAKGOMYQOTRHPFWDGTSCGEGKLCCKGA 503
Db 481 QCRFDGSGYQTCIAFRTEPECKQLWCS-HPDNPYFCKTKGPPLDGTECAPCKWCFKGH 539
QY 504 CVERHNLNKHVDGSKAKWDPPYGPCSCTCGGVQVLARROCTNPTPANGKYCEGVRYKYR 563
Db 540 CIWKSPEQIYQDGGWSSWTKFGSCRSRSGGVRSRSCNNPSPAYGPRCLGPMFEYQ 599
QY 564 SCNLEPCPSASGKSFREOCEAFNGY-----NHSNRLTLAVAVVPKYSVSPRDKCK 617
Db 600 VCNSEECPTIY--EDFRAQCAKRNYSYVHONAKHS-----WVP-YEPDQDAQKCE 647
QY 618 LICRANGTYFYVLAAPKVVVDGTLCS-PDSTSVCVQCKICACDGNLGSKKRFDKCGVCG 676
Db 648 LICQSDATGDDWFMNQVVDHGTCSRDPYSVCARGCEVPVCGDEKGVSKADKCGVCG 707
QY 677 GDNKCKKYTGTLFTK--PMHGYNFVVAIPAGASSIDIRORGYKGLIGDNDYALKNS-QG 733
Db 708 GDNHCRTVKGTGLGKASKAGALKVQIPAGARHIQI-----EALSKSPHRIYVKNQVVG 762
QY 734 KYLLN--GHFVWVAVERDLVVKGSLRLYSGTGTAVESLQASRPILPLTVEVLSVGKMT 791
Db 763 SFILNPKGK--EATSKTFTAMG-LEWEDAVEDAKESLKTSGPLPEAIALAL----- 811
QY 792 PRVRYFYLPKEPREDKSSHKDPGRPS-----VLHNSVLSL--SNQV--EQPDRPP 840
Db 812 -----PTEGGPRSLAYKYVIHEDLLPLIGSNVLEEMD---T 848
QY 841 ARWVAGSWGPCASCGSLQKRAVDGR----- 867
Db 849 YEWALKSWAPCKACCGGIGQFTYKGCRRRRDHMHVORHLCDHKKPKPIRRRCNQHPCSQ 908
QY 868 -----GSAGQRT-----VPACDAARHPVETQACG-----EPC 894
Db 909 PVWVTEWAGACSRSCGKLGVTQIRG IQCLPLSNGTHKVPKACACADGDRPEARPCLRVPC 968
QY 895 PT-WELSAWSPCKSCSGRGFRKSLK-----VHGGRLLRADQCNLHRKPOELDFCVL 947
Db 969 PAOWRLGAWSQCSATCGEGIQOQVYVVCRTNANSLGH-----CEGDR-PDTVQVCSL 1018
QY 948 RPC 950
Db 1019 PAC 1021
RESULT 12
Q8TEY8 PRELIMINARY; PRT; 1159 AA.
ID Q8TEY8
AC Q8TEY8
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

```
DE ADAMTS14.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21839041; PubMed=11741898;
RA Collige A., Vandenbergh I., Thiry M., Lambert C.A., Van Beeumen J.,
RA Li S.W., Prockop D.J., Lapierre C.M., Nussgens B.V.;
RT "Cloning and characterization of ADAMTS-14, a Novel ADAMTS Displaying
RL High Homology with ADAMTS-2 and ADAMTS-3.";
RL J. Biol. Chem. 277:5756-5766(2002).
DR EMBL; AF366351; RAL79814.1;
SQ SEQUENCE 1159 AA; 127336 MW; A5B130149BF7F34 CRC64;

Query Match 21.7%; Score 1120.5; DB 4; Length 1159;
Best Local Similarity 30.0%; Pred. No. 6.2e-85;
Matches 307; Conservative 123; Mismatches 380; Indels 213; Gaps 41;

QY 45 PEDSGDGLIFQITAFQEDFYHLTPDAQFLAPAFSTEH-----LGVPLQGLTGSSD 97
DB 31 PGRVGRHSLYFNVTGKELHLRLRPNRLVPGSSVWEQEDFRELFRQLR----- 82
QY 98 LRCFTSGDVNAEDPSFAVLCGLGURGAFYRGAEYVISPPLNASAPAAQNSQG-AHL 156
DB 83 -QECVVTGGVTGMPGAATAISNCDGLAGLIRTDSTDFIEPLRGO--QKEASGRTHV 138
QY 157 LQRRGVP---GPGSDPTSRCGVASGNWPAIRLALDPYKPRAGFESRRSRGRKR 212
DB 139 YVREAVQOEAEPPDGLHNE-AFGLDLPNLGLVGDO---LGDTERKRHRAPGS 191
QY 213 FVSIPRYVETLVVADESVMKPHGAD-LEHYLLTLTAARLYHPSILNPINIVVVKLL 271
DB 192 Y-----SIEVLLVDDSVVRFGKEHVQNYVLTLMNIVDEIYDESLGVHINIALVRLIM 246
QY 272 LRDRDSGPKV-TGNAALTLRNFCAWQKLNKYSKHPYWDTPILFTRODLGAGTCDTL 330
DB 247 VGYRQSLIERGNPSRSLQVCRWAHSQORQDPSSHAHHDHVFLTRQDF-GPSGMQ-- 303
QY 331 GMADVTGMDCKRSCSVIEDDGLPSAFTTAHELGHVFNPHD---NVKCEEVFKLRAN 387
DB 304 GYAPVTGMCUPLRSCALNIEDGESSAFVIAHETGHVGMHEHVGNGCADETSLGS---- 359
QY 388 HMSNPTLIQIDRANPWSACSAIITDLDLSDGDCLLDQPSKPI-SLPEDLPFCASYTL 446
DB 360 -VMAPLVQAAAFHFRHSRCSKLELSRYLPS--YDCLDDPDPAPWQPPPELPINYSMDE 416
QY 447 QCELAFGVGSKPC---PYWQYCTKLINCTGKAKQWYQTRHPFPWADGTSCGEGKLC 503
DB 417 QCRFDGSGYQTCIAFTFTPECPKQWCS-HPDNPYFCKTKGPPLDGTETCAFGKWC 475
QY 504 CVERHNLNKHVRDGNWAKMDPQPCSRCTGGGVQOLARRQCTNPTPANGKCYCEGVK 563
DB 476 CIWKSPEQTFQDGGHSSWTKFQCSRSCGGGVRSRSCNPNPSPAYGRCPLGPMFEX 535
QY 564 SCNLEPCPSSAGSKSFREEQCAFNGY-----NHSTNRLTLAVAWPKYGVSPRDK 617
DB 536 VCNSECPGTY--EDFRAQCAKRNYSYVHQNAKHS-----WVP-YEPDDDAQKCE 583
QY 618 LICRANGTYFYVLAPKVVVDGTLCS-PDSTSVCQCKCIKACGDGNLGSKKRFDKCG 676
DB 584 LICOSADTGVDVPMNOVHDGTRCSDRDYPSVCARGCEVPVGCDEKVGSMKADKCG 643
QY 677 GDNKSKKVTGLTK--PHGYFNVAIPAGASSIDIRQYKGLTGDNDYLNALNKS-OG 733
DB 644 GDNSHCETVKTGLKASKQAGALKVQIPAGARHIQI-----EALKSPHRIVVKQ 698
QY 734 KYLLN--GHEFVSAVERDLVVKSGLLRYSGTGTAVESLQASRPILEPLTVESLV 791
DB 699 SFILNPKGK---EATSRFTTAMG-LEWEDAVEAKESLKTSGPLPAIALA----- 747
```

```
QY 792 PRVRYSTLPKPEPREDKSSHPKDPGRPS-----VLHNSVLSL--SNQV--EOPDDRPP 840
DB 748 -----PPTGGPRSSLAYKYVIHEDLLPLIGSNVLLLEMD---T 784
QY 841 ARWVAGSWGPCASCSGSLQKRAVDCR----- 867
DB 785 YEWALKSWAPCSKACGGGIGFTYGCRRRRDDHMHVQRHLCDHKRKPPIRRRCNQHP 844
QY 868 -----GSAGQRT-----VPACDAARHPVETQACG-----EPC 894
DB 845 PVWVTEWAGACSRSCGKLGVTQTRGICQCLLPNSNGTHKVPKACAGDRPEARPC 904
QY 895 PT-WELSAWSPCKSCSGRGQRRLSKC-----VHGGRLLARDQCNLHKKPQELDF 947
DB 905 PAQWRLGAWSCSACGEGIQQOVVCRTNANSLSG-----CEGDR-PDTVQVCSL 954
QY 948 RPC 950
DB 955 PAC 957

RESULT 13
Q8TE55 PRELIMINARY; PRT; 1223 AA.
AC Q8TE55;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Metalloprotease-disintegrin protease.
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=FETAL LUNG;
RC MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002);943.1;
DR EMBL; AJ345098; CAC87943.1;
KW Integrin; Protease.
SQ SEQUENCE 1223 AA; 133925 MW; D585B6593977ED15 CRC64;

Query Match 21.7%; Score 1118; DB 4; Length 1223;
Best Local Similarity 29.8%; Pred. No. 1.1e-84;
Matches 305; Conservative 123; Mismatches 379; Indels 216; Gaps 40;

QY 45 PEDSGDGLIFQITAFQEDFYHLTPDAQFLAPAFSTEH-----LGVPLQGLTGSSD 97
DB 98 PGRVGRHSLYFNVTGKELHLRLRPNRLVPGSSVWEQEDFRELFRQLR----- 149
QY 98 LRCFTSGDVNAEDPSFAVLCGLGURGAFYRGAEYVISPPLNASAPAAQNSQG-AHL 156
DB 150 -QECVVTGGVTGMPGAATAISNCDGLAGLIRTDSTDFIEPLRGO--QKEASGRTHV 205
QY 157 LQRRGVP---GPGSDPTSRCGVASGNWPAIRLALDPYKPRAGFESRRSRGRKR 212
DB 206 YVREAVQOEAEPPDGLHNE-AFGLDLPNLGLVGDO-----LGDTERKRHRAPGS 258
QY 213 FVSIPRYVETLVVADESVMKPHGAD-LEHYLLTLTAARLYHPSILNPINIVVVKLL 271
DB 259 Y-----SIEVLLVDDSVVRFGKEHVQNYVLTLMNIVDEIYHDSGLVHINIALVRLIM 313
QY 272 LRDRDSGPKV-TGNAALTLRNFCAWQKLNKYSKHPYWDTPILFTRODLGAGTCDTL 330
DB 314 VGYRQSLIERGNPSRSLQVCRWAHSQORQDPSSHAHHDHVFLTRQDF-----GPS 367
QY 331 GMADVTGMDCKRSCSVIEDDGLPSAFTTAHELGHVFNPHD---NVKCEEVFKLRAN. 387
```



```
Db 368 GVAPVTGCHPLRSCALNHEDGFSAFVIAHETGHVLGMHDGCGCADETSLGS----- 423
Qy 388 HMSPTLIQIDRANPWSAIAITDFLDLSGHGCDLDDPSKPI-SLPEDLPASVYLSQ 446
Db 424 -VMAPLQAAFRHFWSCSLSRPLS--YCLDDPDPAMPQFPPELPLGINSME 480
Qy 447 QCELAFGVSGKPC---PYMQYCTKLWCTGKAKGQMVQCTRHFPWADGTCGEGKLCIKGA 503
Db 481 QCRDFGSGYQCLAFRTFEPCKLWCS-HPDNPFYCKTKRGLPPLDTECAPKWCFKGH 539
Qy 504 CVERHNKLRVDSWAKWDYGPCSTCGGGVGLARQCTNTPANGKYCEGVVYKR 563
Db 540 CIWKSPEQTYQDGGWSWTKFGSCRSRSGGGVRSRSCNPNPAYGGRPCIGPMEYQ 599
Qy 564 SCNLEPCPSSAGSKSFRECEAFNGY-----NHSTNRLTLAVAVWPYKSGVSPRDKCK 617
Db 600 VCNSEECPTG--EDFRAQCAKANSYVYQNAKHS-----WYP-YEPDDDAQKCE 647
Qy 618 LICRANGTYGVYVLPKVVDTGLCS-PDSTSVCVQKCIKAGCDGNLGSKKRFDKCGVCG 676
Db 648 LICQADTGDYVFMNVVHDTRCSYRDPYSVCARGCEVPVGCDEKVEGSKMADKCGVCG 707
Qy 677 GDNCKKVTGLFTK--PMHGYNFVVAIPAGASSIDTRQGYKGLIGDNYLAKNS-QG 733
Db 708 GDNHCHRTVKTGLKSKQAGALKVQIPAGARHIQI-----EALSKSPHRIVVKNQVTG 762
Qy 734 KYLLN--GHFVVSVERDLVVKGLLYSGTGTAVESLOASRPTLEPLTVEVLSVGKMT 791
Db 763 SFILNPKG---EATSTFTAMG--LEWEDAVEDAKESLKTSGPLPEAIAL-- 811
Qy 792 PRVYSYLPKPEPRDKSSHPKDRGPS-----VLHNSVLSL--SNQV--EQPDPRPP 840
Db 812 -----PTEGGPSRLAYKYVIHEDLLPLIGSNVLLLEMD--T 848
Qy 841 ARWAGSWGPCSCSGSLQKRAVDCR----- 867
Db 849 YEWALKSWAPCSKACGGGIRFTYGYCRRRHHVHRHLDCHHKRPKIPRRSNQHPCSQ 908
Qy 868 -----GSAGORT-----VPACDAARHPVETOACG-----EPC 894
Db 909 PVWVTEEWGACRSYKGLGVQTRIGIQCLLPLSNTHKVPAPKACAGRPEARRPCLRVPC 968
Qy 895 PT-WELSAWSPCSKSGRGQRSLK-----VHGGRLLARDQCNLHKPQELDFCVL 947
Db 969 PAOWRLGAWSCSATCGEGIQQRVVCRTNANSILGH-----CEGDR-PDTVQVCSL 1018
Qy 948 RPC 950
Db 1019 PAC 1021
```

RESULT 14

```
Q96L37 ID Q96L37 PRELIMINARY; PRT; 1427 AA.
AC Q96L37;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Von Willebrand factor-cleaving protease precursor.
GN ADAMTS13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX PubMed=11557746;
RA Zheng X., Chung D., Takayama T.K., Majerus E.M., Sadler J.E.,
RA Fujikawa K.;
RT "Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a
RT Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura";
```

```
J. Biol. Chem. 276:41059-41063(2001).
RL EMBL: AY055376; AAL17652.1;
DR MEROPS; M12.241;
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000130; Zn_MTPeptidse.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 4.
DR PROSITE; PS0215; ADAM_MEPPO; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW SIGNAL
FT CHAIN 75 1427 VON WILLEBRAND FACTOR-CLEAVING PROTEASE.
SQ SEQUENCE 1427 AA; 153632 MW; EB18C3AABCI4442 CRC64;

Query Match 17.3%; Score 893.5; DB 4; Length 1427;
Best Local Similarity 29.2%; Pred. No. 1e-65;
Matches 250; Conservative 108; Mismatches 347; Indels 151; Gaps 34;

Qy 174 CGVASQ-WNP-----AILRALDPY-----KPRRAGGESRRSRGRKRF 213
Db 21 CGFLGCGWPSHFQSCIQALEPQAVSSYLSPPGAPLKPSPSPGFQQRQOR-----RA 75
Qy 214 VSTPRYVETLVVADESMTKFGADLEHYLLTLATAARLYRHPSTILNPINIVVVKVLLR 273
Db 76 AGGILHLELLVAVGPDVFAHQEDTERVLTNLTGAELLRDPISLGAOFVRLVRLVLT 135
Qy 274 DRDSGKVTGNAALTLRNFCWOKLKYSDKHPEYWDATILFTRODL-CGATTCDTLGM 332
Db 136 EPEGAPNITANLTSSLLSVCGSQINPEDDTPGHADLVLYITRFDLELDPGNRQVRG 195
Qy 333 ADVGIMCDPKRSCSVIEDDGLPSATTAHELGHVFMNPHDNVVKCEEVFGKRAHNMSP 392
Db 196 TOLGGACSPWTSCLITDGTDFLGVITIAHEIGHSFLEHGDGAPG-----CGPFGHVMAS 251
Qy 393 TLIQIDRNP-----WSACSAIITDFLDSGHGDCLLD-----OPSKPISLPEDLPASVT 443
Db 252 -----DGAAPRAGLAWSPCSRRLQLLSLSAGRARCVWDPPRPPQGSAGHPDPAQCLYIS 306
Qy 444 LSQCELAFVGSKPCPY-----MOYCTKLWCTGKAKGMVQCTRHFPWADGTCGEGKLC 499
Db 307 ANEQCRVAFPGKAVACTFAREHLDMCQALSCHTDPLDQSSCSRLVPLLDGTECGVENKC 366
Qy 500 LKAGACVERHNLNK-HRVDSWAKWDYGPCSTCGGGVGLARQCTNTPANGKYCEGV 558
Db 367 SKGRCSLVELTPIAAVHGRWSSWGPSPCSKSGGGVVTTRRQCNRPAPFGGRACVA 426
Qy 559 RYKRSCLNLEPCPSSAGSKSFRECEAFNG-----YNHSTNRLTLAVAW---V 604
Db 427 DLQAEKMTQACEKQ--LEFMSQOCARTDQGPLRSSPGASFYH-----WGA 474
Qy 605 PRYSGVSPDKCKLICRANGTYGVYVLPKVVDTGLCSP-----DST-SVCVCGKCIAG 658
Db 475 PHSQGD-----LCRHMCRAIGESFIMKRGDSFLDGTCPMPSPREDGTLSLCSVSGSRTFG 531
Qy 659 CDGNLGSKKRFDKCGVCGDNKSKKVTGLFT--KPMHGYNFVVAIPAGASSIDTRQRY 716
Db 532 CDGRMDSQVWDRCVQCGDNDSTCSPRKGSFTAGRAREYVFLVTP--NLTSVYI--ANH 588
Qy 717 KGLIGDDNYLALKNSQGYLLNGHEFVVS--AVERDLVVKGSL-LRYSGTGTAVESLOASR 773
Db 589 RPLF---THLAVRIG-GRYVAVAGKMSISPNTTYPSSLLEDGRVEYRVALTEDRLPLEIR 644
Qy 774 ---PILEPLTVEVL-----SVGKMTPPRVYSFYLPKPEPRDKSSHPKDRGPSVLHNSV 825
Db 645 IWGPLEQEDADIQVYRRYGEEYGNLTRPDITFTYFQK----- 681
Qy 826 LSLSNQVEQPDPRPPARWV-AGSWGPCSCSGSLQKRAVDCRGSAGORTVPA--CDAAH 882
Db 682 -----PQAWVVAARVGPSCVSGAGLRWVNYSCLDQARKELVETVQCGSQ 728
Qy 883 R-PVETQACG-EPDPT-WELSAWSPCSKSGRGFRQRSLKCVGHGRL-----ARQCN 934
```

Db 729 OPPANPEACVLEPCPPYNAVDFGCSASCGGLRERVRVCEAGSLLKTLPPARCRAG 788
QY 935 LHRKPQELDFCVLRPC 950
Db 789 AOPPAVALETCPNPQC 804

RESULT 15

O9GL54 PRELIMINARY; PRT; 269 AA.
AC O9GL54;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Aggrecanase-2 (Fragment).
GN ADAMTS-11.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Goad D.L.; Goad M.E.;
RT "Molecular cloning of lapine aggrecanase-2 (ADAMTS-11) from articular chondrocytes."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBAJ databases.
DR EMBL; AF317415; AAG33062.1; -.
DR HSSP; O9PW35; 1BUD.
DR MEROPS; M12.225; -.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 1.
DR SMART; SM00209; TSPI; 1.
DR PROSITE; PS50215; ADAM_WEPRO; 1.
DR PROSITE; PS50092; TSPI; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 269 269
SQ SEQUENCE 269 AA; 29193 MW; 97A1CA80B33452FA CRC64;

Query Match 16.1%; Score 830; DB 6; Length 269;
Best Local Similarity 54.7%; Pred. No. 1.8e-61;
Matches 146; Conservative 37; Mismatches 74; Indels 10; Gaps 3;

QY 329 TLGMADVGTMCDPKRSCTVEDDGLPSAFTTAHELGHVENMPHDNVKVCVEVFGKLRANH 388
Db 1 TLGMADVGTICSPERSCAVIEDDGLHAFTVAHEIGHLLGLSHDSSKFCENFGSTEDKR 60
QY 389 MMSPTLIQIDRANPWSACSAAIITDPLDSHGDCLLDQPSKIPSLPDLPGASYTLISQOC 448
Db 61 LMSILTSIDASKPWSKCTSAITTEFLDDHGNCLLDVPKQILGPEELPGOTYDATQOC 120
QY 449 ELAFGVGSPCPYMQYCTKLWCTGKAKGQMVQOTRHPWADGTSCEGKCLKGACVER- 507
Db 121 NLTFGPEYTVCPGMDVCARLWCAVVRQGMVCLTKKLPAVEGTPCGKGRICLQKCVDKT 180
QY 508 ----HNLNHRVDGSKAKWDPTGCSRTCTGGVGLARROCTNPTPANGKGYCEGVRYK 563
Db 181 KKKYISTSSH---GNWGSWGPWGQCSRCGGVQFAIRHCNNPAPRNSGRYCTGKRIYR 237
QY 564 SCNLEPCPSSASGKSPFEQCEAFNGY 590
Db 238 SCSVTPCP--ANGKSPRHEQCEAKNGY 262

Search completed: May 9, 2003, 15:23:35
Job time : 54 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:23:41 ; Search time 56 Seconds
(without alignments)
1561.150 Million cell updates/sec

Title: US-09-965-631-4
Perfect score: 5162
Sequence: 1 MLLGLITLAFAGRTAGGSE.....DQCNLHRKPOELDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5162	100.0	950	10	US-09-965-631-4
2	4248.5	82.3	823	9	US-10-163-516-2
3	2486.5	48.2	967	12	US-10-105-929-2
4	2485.5	48.1	950	10	US-09-321-987B-4
5	2485.5	48.1	968	9	US-10-163-316-7
6	2274	44.1	727	9	US-10-097-597-1
7	2274	44.1	727	9	US-10-097-580-1
8	2274	44.1	727	10	US-09-445-023A-1
9	2273	44.0	727	9	US-10-097-597-12
10	2273	44.0	727	9	US-10-097-580-12
11	2273	44.0	727	10	US-09-445-023A-12
12	2115	41.0	905	10	US-09-918-171A-9
13	1965	38.1	1629	10	US-09-972-467-2
14	1919	37.2	367	10	US-09-965-631-6
15	1914	37.1	837	9	US-10-174-590-352
16	1914	37.1	837	9	US-10-176-758-352
17	1914	37.1	837	9	US-10-175-737-352
18	1914	37.1	837	9	US-10-173-706-352
19	1914	37.1	837	9	US-10-175-738-352

ALIGNMENTS

* RESULT 1
US-09-965-631-4
; Sequence 4, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020115842A1el Human Proteases and Polypeptides Enco
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-965-631-4

Query Match 100.0%; Score 5162; DB 10; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLLGLITLAFAGRTAGGSEPERVVVRLDPDINGRRYYWRGSDSDGGLIFQITAF	60
Db	1	MLLGLITLAFAGRTAGGSEPERVVVRLDPDINGRRYYWRGSDSDGGLIFQITAF	60
QY	61	QEDFYHLTPDQAFLAPAFSTHGLGVPLQGLTGGSSDLRCRCFYSGDVNAEPDSFAVSLC	120
Db	61	QEDFYHLTPDQAFLAPAFSTHGLGVPLQGLTGGSSDLRCRCFYSGDVNAEPDSFAVSLC	120
QY	121	GGIRGAFGYRGAEEYVTSPLPNASAPAAQRNSOGAHLRRGYPGGPSGDPTRSCVSGW	180
Db	121	GGIRGAFGYRGAEEYVTSPLPNASAPAAQRNSOGAHLRRGYPGGPSGDPTRSCVSGW	180
QY	181	NPAILRALDPYKPRAGGESRRRSRAKRFVSIPIRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRAGGESRRRSRAKRFVSIPIRYVETLVVADESMVKFHGADLEH	240
QY	241	YLLTLLATAARYRHPSTILNPINIVVVKLLLRDSDSGPKVTGNAALTLRNFCAMQKLN	300

Db 241 YLLTLLTAARLYRHSILNPIVVKVLLLRDSDGPKYGNAAFLNFCWAKKLN 300
QY 301 KVSXKHPEYWDTAILFTRODLGATTCDTLGMADVGTWCDPKRSCSVIECDGLPSAFTTA 360
Db 301 KVSXKHPEYWDTAILFTRODLGATTCDTLGMADVGTWCDPKRSCSVIECDGLPSAFTTA 360
QY 361 HELGHVFNPHDNVKEVFCGLKLANHMSPTLTIIDRANPWSACSAAITDFLDSGHG 420
Db 361 HELGHVFNPHDNVKEVFCGLKLANHMSPTLTIIDRANPWSACSAAITDFLDSGHG 420
QY 421 DCLLDQSPKIPISLPEDLPASVYTLSSQCELAFLGVSKPCPYMOCYCKLWCTGKAKGQWVC 480
Db 421 DCLLDQSPKIPISLPEDLPASVYTLSSQCELAFLGVSKPCPYMOCYCKLWCTGKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYDGVGSAKWADPYGPCSETCGGVQVLAR 540
Db 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYDGVGSAKWADPYGPCSETCGGVQVLAR 540
QY 541 ROCTNPTPANGGKYGCVGRVYKRSNLEPCPSSASGKSFREOQCEAFNGYNHSTNRLTLA 600
Db 541 ROCTNPTPANGGKYGCVGRVYKRSNLEPCPSSASGKSFREOQCEAFNGYNHSTNRLTLA 600
QY 601 VAWVPYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGFTLCSPDSTSVCVQCKIKAGCD 660
Db 601 VAWVPYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGFTLCSPDSTSVCVQCKIKAGCD 660
QY 661 GNLSKRRDKCGVCGDGNKSKKVTGLTKPMHGFYFVAIPAGASSIDIRQGYKGLI 720
Db 661 GNLSKRRDKCGVCGDGNKSKKVTGLTKPMHGFYFVAIPAGASSIDIRQGYKGLI 720
QY 721 GDDNYLALNKSQGYLLNGHFVYSAVERDLVWKSLLRYSGTGTAVESLQASRPILPLT 780
Db 721 GDDNYLALNKSQGYLLNGHFVYSAVERDLVWKSLLRYSGTGTAVESLQASRPILPLT 780
QY 781 VEVLVSGKMTPPRVYSFYLPKPEPRDKSSHPRGDSVLSNLSNOVEQPDPRPP 840
Db 781 VEVLVSGKMTPPRVYSFYLPKPEPRDKSSHPRGDSVLSNLSNOVEQPDPRPP 840
QY 841 ARWVAGSWGPCASCSGGLQKRAVDCRGSGAGORTVPACDAHRPVEVTOACGPECPITWELS 900
Db 841 ARWVAGSWGPCASCSGGLQKRAVDCRGSGAGORTVPACDAHRPVEVTOACGPECPITWELS 900
QY 901 AWSPCSKSGRGFRRLSKCVGHGRLARQDQCNLHRKPJELDFCVLRPC 950
Db 901 AWSPCSKSGRGFRRLSKCVGHGRLARQDQCNLHRKPJELDFCVLRPC 950

RESULT 2
US-10-163-316-2
; Sequence 2, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MP101-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-316-2
Query Match 82.3%; Score 4248.5; DB 9; Length 823;
Best Local Similarity 97.1%; Pred. No. 6.9e-295;
Matches 792; Conservative 0; Mismatches 1; Indels 23; Gaps 1;

QY 1 MLLGLILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
Db 1 MLLGLILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
QY 61 QEDFYHLTTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
Db 61 QEDFYHLTTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
QY 121 GGLRGAFYGAEVISPLPNASAPAAORNSOGAHLQRRVPGPGSGDPTSCGVSASGW 180
Db 121 GGLRGAFYGAEVISPLPNASAPAAORNSOGAHLQRRVPGPGSGDPTSCGVSASGW 180
QY 181 NPAILRALDYPKRRAGFEGESRRRSRRAKRFVSPRYVETLWVADESMVKFHGADI EH 240
Db 181 NPAILRALDYPKRRAGFEGESRRRSRRAKRFVSPRYVETLWVADESMVKFHGADI EH 240
QY 241 YLLTLLTAARLYRHSILNPIVVKVLLLRDSDGPKYGNAAFLNFCWAKKLN 300
Db 241 YLLTLLTAARLYRHSILNPIVVKVLLLRDSDGPKYGNAAFLNFCWAKKLN 300
QY 301 KVSXKHPEYWDTAILFTRODLGATTCDTLGMADVGTWCDPKRSCSVIEDDGLPSAFTTA 360
Db 301 KVSXKHPEYWDTAILFTRODLGATTCDTLGMADVGTWCDPKRSCSVIEDDGLPSAFTTA 360
QY 361 HELGHVFNPHDNVKEVFCGLKLANHMSPTLTIIDRANPWSACSAAITDFLDSGHG 420
Db 361 HELGHVFNPHDNVKEVFCGLKLANHMSPTLTIIDRANPWSACSAAITDFLDSGHG 420
QY 421 DCLLDQSPKIPISLPEDLPASVYTLSSQCELAFLGVSKPCPYMOCYCKLWCTGKAKGQWVC 480
Db 421 DCLLDQSPKIPISLPEDLPASVYTLSSQCELAFLGVSKPCPYMOCYCKLWCTGKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYDGVGSAKWADPYGPCSETCGGVQVLAR 540
Db 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRYDGVGSAKWADPYGPCSETCGGVQVLAR 540
QY 518 SWAKWDPYGPCSRTCGGVQVLARQCTNPPANGKKYCEGVRYKRSNLEPCPSSASGK 577
Db 541 SWAKWDPYGPCSRTCGGVQVLARQCTNPPANGKKYCEGVRYKRSNLEPCPSSASGK 600
QY 578 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCKLICRANGTGYFYVLAPKVV 637
Db 601 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCKLICRANGTGYFYVLAPKVV 660
QY 638 GTLCSPOSTSVCVQCKIKAGCDNGLSKRFRDCKGCGVDNKSCKKVTGLTKPMHGIN 697
Db 661 GTLCSPOSTSVCVQCKIKAGCDNGLSKRFRDCKGCGVDNKSCKKVTGLTKPMHGIN 720
QY 698 FVWAIAPAGASSIDIRQGYKGLIGDDNYLALNKSQGYLLNGHFVWSAVERDLVWKSLL 757
Db 721 FVWAIAPAGASSIDIRQGYKGLIGDDNYLALNKSQGYLLNGHFVWSAVERDLVWKSLL 780
QY 758 RYSGTGTAVESLQASRPILPLTVEVLVSGKMTPPR 793
Db 781 RYSGTGTAVESLQASRPILPLTVEVLVSGKMTPPR 816
RESULT 3
US-10-105-929-2
; Sequence 2, Application US/10105929
; Patent No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961

;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2
;; LENGTH: 967
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-105-929-2

Query Match 48.2%; Score 2486.5; DB 12; Length 967;
Best Local Similarity 49.0%; Pred. No. 4.5e-169;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;
QY 1 MLLLGILTLAFAGTAGSEPEREVVPIRLDPDINGRRYVRGPEDESGDGLQIFQITAF 60
Db LLLLLAALLAVSDALGRSEDEELVLP-ELE-----RAP---GHGTTIRLRHAF 81
QY 61 QDEFYHLTPDPAQFLAPAFSTEHLG-----VPLGGLGSGSDLRRCFYSGDVNAEPDSF 114
Db DQQLDLELRPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFYGAAYVVISPLPNAS---APAAQNSOGA-----HLLQ---RRCVPGG 165
Db AALSUCGVRGAFYLLGGEAYFIQPLPAASERLATAAGKEKPPAPLQFHLRRNQGDYGG 196
QY 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRRAGFSGSRRRS 207
Db TCGVDDDEPRPTGKAETDEDEGTGEDEGPQWS-----PQDPALQGVGP-TGTGS 247
QY 208 GRAKRFVSPRYVETLVVADESMVKFHGADLEHYLLTLATAARLYRHPS:LNFINIVVV 267
Db IRKKRFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFVAARLYKHPS-RNSVSLVV 307
QY 268 KYLLLRDRSGKVTGNALTLRNCANQKLNKVDKHPKWDYTAIFYHQDLGGATTC 327
Db KILVTHDEQKGPVTSNAALTLRNCQKHNQKHNQKHNQKHNQKHNQKHNQKHNQKHNQKHN 367
QY 328 DTLGMADVGTMDCKPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKNVCEVEFKLRAN 387
Db DTLGMADVGTMDCKPRKSCSVIEDDGLQAAFTTAHELGHVFNPHDNNVKNVCEVEFKLRAN 427
QY 388 HMSPTLIQIDRANPWSACSAIITDLDGSHGDCLLDPKSKPISLPDLPGASYTLQQ 447
Db HMAWSMLSLNDSQWSPSCSAAYMTSFLDNHGECLMDKPNQPIQLPGDLPGTSDANRQ 487
QY 448 CELAFGVGSKPCP-YMQYCTKLWCTGKAKQWVCOTRHPFADGTSCGEGKLCUKGACVE 506
Db COFTGEDSKHCPDAASTCTLWCTGTSGGLVLCQTKHFPWADGTSCGEGKWCINGKCVN 547
QY 507 RHNLNKH---RYDGSWAKWDYPGCSRTCGGGVQLARRQCTNPTPANGSKYCEGVYKVR 563
Db KTD-RKHETPPHGSWGMWPGWDCSRTCGGGVQVYTMRECONPVKNGKKEGKRVYR 606
QY 564 SCNLEPCPSASGSKFREEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRDKCLICRAN 623
Db SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAWEPYKAVGSPKDRCKLICQAK 665
QY 624 GTCGYFVLAPKVDVDTGLSPDSTSVCGVQKCIKAGCDGNLGSKRFKQVCGGDNKCK 683
Db GIGYFVLQPKYVDVDTGTPCSDSTSVCGVQKCIKAGCDRIIDSKKKFKDCGVCVGGNGSTCK 725
QY 684 KYTGFTKPMHGYVFWAIPAGASSIDIRORYKGLIGDNDYALKNSQKYLNLNGHVVV 743
Db KISGVSATKPGYHDIITPTGATNIEVQRNQRNNGSLAIIKAADGYIILNGDYTL 785
QY 744 SAVERDLVVKGLSYSGTGTAVESLOASRPILEPTVEVLSVGKWTPPRVRYSFYLPKE 803
Db STLEQDVMYKGVVLYSGSSAALERIRSFSLPKEPLTIQVLTGNALRPKIKYIFYVKK 845
QY 804 PREDKSSHPKDRGRFSVLHNSVLSNQVEQPDPRPARWAGSWGPCASCGGSLQKRA 863
Db ----RES-----FNAIPTS-----AWVTEWGECSKCELGWORRL 878

QY 864 VDCRSAGQRTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCKSCGGRGFRRSIKC 920
Db 879 VECEDINGQ---PASECAKEVKPASTPFCADHCFQWQLGWSCKTCKGKGYKKRSIKC 935
QY 921 VGHGRLRLARDQCNLHRRKPOE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKPKRHFIDFCTMAEC 966
RESULT 4
US-09-321-987B-4
; Sequence 4, Application US/09321987B
; Patent No. US20020102210A1
; GENERAL INFORMATION:
; APPLICANT: Kimble, Judith E
; APPLICANT: Blueloch, Robert H
; TITLE OF INVENTION: Agent and Method for Modulating Cell Migration
; FILE REFERENCE: 960296.95386
; CURRENT APPLICATION NUMBER: US/09/321,987B
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087,170
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/129,023
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Murine
US-09-321-987B-4

Query Match 48.1%; Score 2485.5; DB 10; Length 950;
Best Local Similarity 48.0%; Pred. No. 5.2e-169;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;
QY 1 MLLLGILTLAFAGTAGG--SEPEREVVPIRLDPDINGRRYVRGP-EDSGQGLIFQI 57
Db 20 LLLASITMLLCARGHGRTEDEELVLP-SLE-----RAPGHDSTIRL--RL 66
QY 58 TAFQEDYHLTPDPAQFLAPAFSTEHLGVPLOGLTGG-----SDLRRCYSGDVNA 109
Db DAFQQLHLKLPQDSGLAPGFTLQTV-----GRSPGSEAQLHDPDGLAHCFYSGTVNG 121
QY 110 EPDSFAVSLCGGLRGAFYGAAYVVISPLPNAS---APAAQNSOGA-----HLLQRRG 161
Db DPGSAALSLCEGVRGAFYLOGEEFFIQAPAGVATERLAPAVPEESSARPQHILRRR- 180
QY 162 VPGGPGDPTSRCSVASGWNPAILRALDPYKPRRAGFSGSRRR----- 206
Db 181 ----RRSGGAKCGVMD-----DETLPTSDSPESQNTNQPVRDPTPDAGKP 226
QY 207 ----SGRAKRFVSPRYVETLVVADESMVKFHGADLEHYLLTLATAARLYRHPSILNFI 262
Db SGPSIRKKRFVSPRYVETMLVADQSMADFHGSLGKHYLLTLFVAARLYKHPSIRNFI 286
QY 263 NIWVVKVLLLRDRSGKVTGNALTLRNCANQKLNKVDKHPKWDYTAIFYHQDLG 322
Db SLVVVKVLTVEEQKGPVTSNAALTLRNCQKHNQKHNQKHNQKHNQKHNQKHNQKHNQKHN 346
QY 323 GATTCDFTLGMADVGTMDCKPRKSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKNVCEVEFG 382
Db GSHTCDFTLGMADVGTMDCKPRKSCSVIEDDGLQAAFTTAHELGHVFNPHDNNVKNVCEVEFG 406
QY 383 KLRANHMSPTLIQIDRANPWSACSAIITDLDGSHGDCLLDPKSKPISLPDLPGASY 442
Db VTGDSHLSMWSLSDHDSQWSPSCSAAYMTSFLDNHGECLMDKPNQPIQLPGDLPGT 466
QY 443 TLSQCELAFGVSKPCP-YMQYCTKLWCTGKAKQWVCOTRHPFADGTSCGEGKLCUK 501
Db DANRCQFTFEESKHCPCDAASTCTLWCTGTSGGLVLCQTKHFPWADGTSCGEGKWCVS 526
QY 502 GACVVERHNLNKH---RYDGSWAKWDYPGCSRTCGGGVQLARRQCTNPTPANGSKYCEGV 558

```
Db 527 GKCVNKTDM-KHEATPVHSGSWGFWGDCSRFCGCGVYTHRECDNPVPKNGGKYCEGK 585
QY 559 RVKYRSCNLEPCPSSAGSKSFRFEEQCEAFNGYNHSTNRLTLAVWVPKYSGVSPDRCKKL 618
Db 586 RVYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTWETPKYAGVSPDRCKL 644
QY 619 ICRANGTGYFVLAPKVVVDGTLCSPDSTSVCGVGKICAKGCDNLGSKKRFKCGVCGGD 678
Db 645 TCEAKGIGYFVLQPKVVVDGTPCSPDSTSVCGVGQCVKAGCDBTIDSKKKFKCGVCGGN 704
QY 679 NKSCKVTGLTKPMHGYNFVVAIPAGASSIDIRQGYKGLICDDNYLALKNSQGYLLN 738
Db 705 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHNRQGRNNGSFLAIRAADGTIYL 764
QY 739 GHFVSAVERDLVVKGLSLRYSGTGTAVESLQASRIEPLTVEVLSVGKNTPPRVYSF 798
Db 765 GNFTLSTLEQDLYKTGLVRLYSGSSAALIRSFSLKPEPLTIQVLMVGHALRPKIKETY 824
QY 799 YLPKPREDKSSHPKDPGRPSVLHNSVLSNOVEQDPPRPAWVAGSNGPCSCSGSG 858
Db 825 FMKKKTES-----FNAIPTFS-----FWVIEEWGECSTKCGSG 857
QY 859 LOKRAVDCRGSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSCGRGFOR 915
Db 858 WQRRVWQCRDINGH---PASECAKEVKPASTRPCADLPCPHQVGDWSPCSTCKGKYK 914
QY 916 RSLKCVGHGGRLLARDQCNLHRRKPQ-E-LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSNESCDPLKPKHYIDFCTLTQC 950

RESULT 5
US-10-163-316-7
; Sequence 7, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI01-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-163-316-7
```

```
Query Match 48.1%; Score 2485.5; DB 9; Length 968;
Best Local Similarity 48.0%; Pred. No. 5,4e-169;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

QY 1 MLLGLTILAFAGRTAGG--SPEREVVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI 57
Db 37 LLLASITMLLCARGAHRPTDEDELVP-SLE-----RAPGHDSTTTLR--RL 83
QY 58 TAPQEDFYHLHPDQAQFLAPASTEHLGVPLQGLTGG-----SDLRFCFYSGDVNA 109
Db 84 DAFGQQLHLKLOPDSGLFAPGFTLQTV-----GRSPGSEAOHLDPDGLAHCFYSGTVNG 138
QY 110 EPDSFAVSLCGRLGAFAGYGAEEYVISPLPNAS-----APAAQRNSQGA-----HLLQRRG 161
Db 139 DPGSAALSICGVRGAFYLOGEEFFIQPAGVATERLAPAVEBESSARPQFHLIRRR- 197
QY 162 VPGPGSDPTSCGVASGWNPAIRALDPYKPRRAGFGESRRR----- 206
Db 198 ----RRSGGAKCGVMD-----DETLPSTSDRPESQNRNQWPFVRDPTPDAGKP 243
```

```
QY 207 ----SGRAKRFVSTPRYVETLVVADESMKVFHGGADLEHYLLTLLTAARLYRHPSTIN 262
Db 244 SGPGSKIRKKRFSRPRYVETMLVADQSMADFHGSLGKHYLLTFSVAAREYKHPSPNSI 303
QY 263 NIVVVKVLLLRDRDSGPKVTGNAAALTNRNCAWOKKLINKYSDKHPEYWDTAILFTRCCLC 322
Db 304 SLVVVKLVIIYEEQKGEVTSNAAALTNRNCAWOKKNSPDRDPEHYDTAILFTRODLC 363
QY 323 GATTCDTLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNPMHDPNVKVCVEVFG 382
Db 364 GSHTCDTLGMADVTCVCDPSRSCSVIEDDGLQAFTTAHELGHVFNPMHDPNDAKHCASLNG 423
QY 383 KLRANHMSPTLIOLDIRANPWSACSAAIITDFDSHGDCDLLDQPSKPISLPEDLPGASY 442
Db 424 VTGSHLMASMSLSDHSPWSPCSAYMVTFSFLDNHGHECLMDKPNQPIKLPDLPGLTY 483
QY 443 TLSOCELAFCGVGSKPCP-YMYCTCKLWCTGKAGQWVOCOTRHFPMADGTSCEGKLCCK 501
Db 484 DANQCQPTFGESKHCPCDAASTCTTLWCSTGGLVLCQTKHPPWADGTSCEGKWCVS 543
QY 502 GACYERHNLANKH---RYDGSWAKWDYPGCSRTCGGQVQLARRQCTNPTPANGSKYCEV 558
Db 544 GKCVNKTDM-KHEATPVHSGSWGFWGDCSRFCGCGVYTHRECDNPVPKNGGKYCEGK 602
QY 559 RVKYRSCNLEPCPSSAGSKSFRFEEQCEAFNGYNHSTNRLTLAVWVPKYSGVSPDRCKKL 618
Db 603 RVYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTWETPKYAGVSPDRCKKL 661
QY 619 ICRANGTGYFVLAPKVVVDGTLCSPDSTSVCGVGKICAKGCDNLGSKKRFKCGVCGGD 678
Db 662 TCEAKGIGYFVLQPKVVVDGTPCSPDSTSVCGVGQCVKAGCDBTIDSKKKFKCGVCGGN 721
QY 679 NKSCKVTGLTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGYLLN 738
Db 722 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHNRQGRNNGSFLAIRAADGTIYL 781
QY 739 GHFVSAVERDLVVKGLSLRYSGTGTAVESLQASRIEPLTVEVLSVGKNTPPRVYSF 798
Db 782 GNFTLSTLEQDLYKTGLVRLYSGSSAALIRSFSLKPEPLTIQVLMVGHALRPKIKETY 841
QY 799 YLPKPREDKSSHPKDPGRPSVLHNSVLSNOVEQDPPRPAWVAGSNGPCSCSGSG 858
Db 842 FMKKKTES-----FNAIPTFS-----FWVIEEWGECSTKCGSG 874
QY 859 LOKRAVDCRGSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSCGRGFOR 915
Db 875 WQRRVWQCRDINGH---PASECAKEVKPASTRPCADLPCPHQVGDWSPCSTCKGKYK 931
QY 916 RSLKCVGHGGRLLARDQCNLHRRKPQ-E-LDFCVLRPC 950
Db 932 RTLKCVSHDGGVLSNESCDPLKPKHYIDFCTLTQC 967
```

RESULT 6

```
US-10-097-597-1
; Sequence 1, Application US/10097597
; Publication No. US20030022352A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Ei-ji
; APPLICANT: Hakezaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same.
; CURRENT APPLICATION NUMBER: US/10/097,597
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03
```

```

; PRIOR APPLICATION NUMBER: JP 9-160422
;
; PRIOR FILING DATE: 1997-06-03
;
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
;   LENGTH: 727
;   TYPE: prt
; ORGANISM: Homo sapiens
US-10-097-597-1

```

Query Match 44.1%; Score 2274; DB 9; Length 727;
Best Local Similarity 54.4%; Pred. NO. 4.5e-154;
Matches 408; Conservative 129; Mismatches 173; Indels 40; Gaps 11;

[illegible]

```

? APPLICANT: Inoguchi, Eiji
? APPLICANT: Hakozaiki, Michinori
? APPLICANT: Ishioka, Keiko
? APPLICANT: Ishida, Yukako
? APPLICANT: Matsushima, Kouji
? APPLICANT: Kuno, Kouji
? TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceut
? TITLE OF INVENTION: composition and method of immunologically analyzing human
? FILE REFERENCE: Q57092
? CURRENT APPLICATION NUMBER: US/10/097,580
? PRIOR FILING DATE: 2002-03-15
? PRIOR APPLICATION NUMBER: 09/445,023
? PRIOR FILING DATE: 1999-12-03
? PRIOR APPLICATION NUMBER: JP 9-160422
? PRIOR FILING DATE: 1997-06-03
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 1
? LENGTH: 727
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-10-097-580-1

```

[illegible]

Db 640 ECRDINGQ---PASECAKEVKPASTRPCADHPCPQWQLGEMSSCKTCGKYKRSKLKCL 696
QY 922 GHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 697 SHDGGVLSHESCDPLKKPKHFIDFCTLTQC 726

RESULT 8

US-09-445-023A-1
; Sequence 1, Application US/09445023A
; Patent No. US20020119167A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADAMTS
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-445-023A-1

Query Match 44.1%; Score 2274; DB 10; Length 727;
Best Local Similarity 54.4%; Pred. No. 4,5e-154;
Matches 408; Conservative 129; Mismatches 175; Indels 40; Gaps 11;

QY 209 RAKRFVSIPIRYVETLVVADESVMVKFHGADLEHYLLTLATAARLYRHPSILNFINIVVK 268
Db 9 RKRREVSPIRYVETLVVADESVMVKFHGADLEHYLLTLATAARLYRHPSILNFINIVVK 68
QY 269 VLLLRDRSDGPKVTGNAALTLRNFCAWKKLNKVSDDKHPYWDTAILFTRODLGGATCD 328
Db 69 ILVIIEQKGPVETSNAAALTLRNFCAWKKLNKVSDDKHPYWDTAILFTRODLGGATCD 128
QY 329 TLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVENMHDNVKVCVEVFGKLRANH 388
Db 129 TLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVENMHDNVKVCVEVFGKLRANH 188
QY 389 MMSPTLIQIDRANPWSACSAIITDFLDSDHGDCLLDQPSKPISLPDLPGASYTLSSQOC 448
Db 189 MMSPTLIQIDRANPWSACSAIITDFLDSDHGDCLLDQPSKPISLPDLPGASYTLSSQOC 248
QY 449 ELAFGVSKPCP-YMOCYCTKLWCTGKAKQWVCOTRPFHPWADGTSCEGKILCKGACVER 507
Db 249 QFTFGEDSKHCPDAASTCTTLCWCTGSGGLVLCQTKHFPWADGTSCEGKILCKGACVER 308
QY 508 HNLNKH---RVDGSWAKWDPPYGCSTRTCGGVQLARRQCTNPTPANGKYCEGVVRKYRS 564
Db 309 TD-RKHEDTPFHGSWGPWPGWDCSTRTCGGVQVYTHRECDNPVKNPKYCEGVVRKYRS 367
QY 565 CNLEPCSSASGKSFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPKCKLICRANG 624
Db 368 CNLEPCPDN-NGKTFREOCEAHNEFSKASFGSGPAVEIPKYAGVSPKDRCKLICQAKG 426
QY 625 TGFFYLAKPVVDGTLCSPDSTSVQVQKCIKAGDGNLGSKKRPDKCVCVGDNNKSKCK 684
Db 427 IGYFFVLQPKVDGTPCSPDSTSVQVQKCIKAGDGNLGSKKRPDKCVCVGDNNKSKCK 486
QY 685 VTGLFKPMHGYNFVVAIPAGASSIDIRQYKGLIGDNDYLNALNKSQCKYLINGHFVYS 744
Db 487 ISGSVTSARPGYHDIIVITGATNIEVKQNRQGRNNGSFLAIKAADGTIILNGDYTLS 546

QY 745 AVERDLVVGSLRLRYSGTGTAVESLQASRPILPLAVEVLSVGKMTPPRVRYSFYLPKREP 804
Db 547 TLEQDIMYKGVLRYSGSSAALERIRSFSLKEPLTQVLTGVNLRPKIKYIFYVKKK- 605
QY 805 REDKSSHPKDPGRPSVLHNSVLSNQVEQDDPRPARVWAGSWGSPCSASCGLQKRAV 864
Db 606 ---KES-----FNAITFS-----AWVIEEMGECSSKSELGWORRLV 639
QY 865 DCRGSAGORTVPACDAH--RPVETQACGE-PCPTWELSAWSPCSKSCGGRFORSLKCV 921
Db 840 ECRDINGQ---PASECAKEVKPASTRPCADHPCPQWQLGEMSSCKTCGKYKRSKLKCL 696
QY 922 GHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 697 SHDGGVLSHESCDPLKKPKHFIDFCTLTQC 726

RESULT 9

US-10-097-597-12
; Sequence 12, Application US/10097597
; Publication No. US20030022352A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same,
; TITLE OF INVENTION: pharmaceutical
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADA
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/10/097,597
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-097-597-12

Query Match 44.0%; Score 2273; DB 9; Length 727;
Best Local Similarity 53.7%; Pred. No. 5.4e-154;
Matches 403; Conservative 131; Mismatches 176; Indels 40; Gaps 10;

QY 209 RAKRFVSIPIRYVETLVVADESVMVKFHGADLEHYLLTLATAARLYRHPSILNFINIVVK 268
Db 9 RKRREVSPIRYVETLVVADESVMVKFHGADLEHYLLTLATAARLYRHPSILNFINIVVK 68
QY 269 VLLLRDRSDGPKVTGNAALTLRNFCAWKKLNKVSDDKHPYWDTAILFTRODLGGATCD 328
Db 69 ILVIIEQKGPVETSNAAALTLRNFCAWKKLNKVSDDKHPYWDTAILFTRODLGGATCD 128
QY 329 TLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVENMHDNVKVCVEVFGKLRANH 388
Db 129 TLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVENMHDNVKVCVEVFGKLRANH 188
QY 389 MMSPTLIQIDRANPWSACSAIITDFLDSDHGDCLLDQPSKPISLPDLPGASYTLSSQOC 448
Db 189 MMSPTLIQIDRANPWSACSAIITDFLDSDHGDCLLDQPSKPISLPDLPGASYTLSSQOC 248
QY 449 ELAFGVSKPCP-YMOCYCTKLWCTGKAKQWVCOTRPFHPWADGTSCEGKILCKGACVER 507
Db 249 QFTFGEDSKHCPDAASTCTTLCWCTGSGGLVLCQTKHFPWADGTSCEGKILCKGACVER 308
QY 508 HNLNKH---RVDGSWAKWDPPYGCSTRTCGGVQLARRQCTNPTPANGKYCEGVVRKYRS 564

Best Local Similarity 53.7%; Pred. No. 5.4e-154;
Matches 403; Conservative 131; Mismatches 176; Indels 40; Gaps 10;

```
QY 209 RAKRFVSPRYVETLVADSVKHEGADLEHYLLTLATAARLYRHPSILNINIVVVK 268
Db 9 RKKRFVSSPRYVETMLVADQSDHSGSLKHYLLTFSVAARFYKHPSIRNSISLVVVK 68
QY 269 VLLLRDSDGPKVGTGNAALTFRNCAWOKLNKVDKHEPYWDATLFTRODLGATTC 328
Db 69 ILVIEEQGPEVTNNAALTFRNCSWQKOHNSPSDRPEHYDTAILFTRODLGSHTC 128
QY 329 TLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDMNVKCFEVEFKLRANH 388
Db 129 TLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDMNVKCFEVEFKLRANH 188
QY 389 MMSPTLIQIDRANPWSACSAAITDPLDSHGDCDLDOPSKEPISLPEDLPGASYLTSQC 448
Db 189 LMASMLSSLDHSPQSAVMYTFLDNGHGECLMDKPNPIKLPDLPGLTYDANQDC 248
QY 449 ELAFVGSKPCP-YMOYCTKLMCTGKAKQMVQCTRHFPWADGTSCGEGKICLKAGACVER 507
Db 249 QTFEGESKHCDAASTCTTACTGTSGLLVCQTKHPWADGTSCGEGKICLKAGACVER 308
QY 508 HNLNKH---RVDSNAKWDYGPCSRTCGGGVQLARRQCTNTPANGKYGCEGVKRYRS 564
Db 309 TDM-KHFATPVHSGWGPWGDGSCRTCGGVQVYTMRECDNPVKNGGKYCEGKRYRS 367
QY 565 CNLEPCSSASGSPRECEAFNGYNHSTNRLTLAVAVPVKSGVSPRDKCLICRANG 624
Db 368 CNLEPCSSASGSPRECEAFNGYNHSTNRLTLAVAVPVKSGVSPRDKCLICRANG 426
QY 625 TGYFFVLAPVVDGTLCSPDSTSVQVQKCIKAGCDNGLSKRFDKGVCGGDNKSKCK 684
Db 427 IGYFFVLQPKVVDGTLCSPDSTSVQVQKCIKAGCDNGLSKRFDKGVCGGDNKSKCK 486
QY 685 VTGLTKPMHGNFVVAIPAGASSIDIRQYKGLIGDNNYALKNQKYLNGHFVVS 744
Db 487 MSGIVTSTRPGYHDIITPAGATNIEVKHNRQGRNNGSFLAIRAADGTYLNGNFTLS 546
QY 745 AVERDLVWGLRLLRYSGTGTAVESLQASRPILPTEVYVSGKMTPPRVRYSFYLKREP 804
Db 547 TLQDLYTKGTVLRYSSSAALIRSFPLAKEPLTIQVLMVGHALRPKIKETPFMKKT 606
QY 805 REDKSHPKDRGSPVLHNSVLSNQVEQDDRRPARWAGSWGSPCASCSGLQKRAY 864
Db 607 ES-----FNAITFS-----EWVIEWGECSKTCGSGWQRRVY 639
QY 865 DCRGSAGQRTVPACDAH--RPVETOACGE-PCPTWELSAWSPCKSCGGRGFORSLKCV 921
Db 640 QCRDINGH---PASECAKEVKPASTRPCADLPCPHWQVGDWSPCKSKTCGKGYKKTKCV 696
QY 922 GHGGRLLARDQCNLHRKQOE-LDFCVLRPC 950
Db 697 SHDGGVLSNESCDPLKPKHYIDFCILTQC 726
```

RESULT 12

US-09-918-171A-9
; Sequence 9, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9

; LENGTH: 905
; TYPE: PRT
; ORGANISM: Mus musculus ADAMTS-8
US-09-918-171A-9

Query Match 41.0%; Score 2115; DB 10; Length 905;

Best Local Similarity 45.6%; Pred. No. 1.3e-142;
Matches 431; Conservative 150; Mismatches 259; Indels 106; Gaps 25;

```
QY 2 LLLGILTL----AFAGRTAG---GSEPEREVVPIRLDPDINGRRYYWRGPDSDGQGLI 54
Db 13 LLLLLLQPPPLVCGAPAGPGTGAQAS-ELVVPTRL-----PGSASE--LA 56
QY 55 FOITAFQEDFYDLHTLDAQFLAPAFSTEHLGVPLQGLTGGSSDLRCFYSGDVNABPDSF 114
Db 57 FHLSAFGQGVFLRLAPDASFLAPEFKIERLG-GSSAAGGEPGLRGCFPSGTNGRESL 115
QY 115 AAYSLCGGLRGAGYRGAEYVISPLPNASAPAAQRNSOGA-----HLLORRGVPGGES 167
Db 116 AAMSCVAGWSGSLLAGEEFTIQP-----QGAGDSLDQPHRLQRWG-PCRR 161
QY 168 GDPTSRGCVASGWNPAITLRALDPYKPRRAGFGESES-----RRRSG----- 208
Db 162 EDP----GLAAAEVFPPLQGLEWEVEMNGCQOQERSDNEEDKKQDKREGLLKETEDSRKVP 217
QY 209 -----RAKRFVSPRYVETLVADSVKHEGADLEHYLLTLATAARLYRHPSILNIP 261
Db 218 PPFSGKTRSKRFVSEARFVETLLVADASMAAFYGTDLQNHILTVMSMAARIYKHPISRS 277
QY 262 INIVVVKVLLLRDSDGPKVGTGNAALTFRNCAWOKLNKVDKHEPYWDATLFTRODL 321
Db 278 VNLVVVVKVLEVERKWEVSDNGGLFLRNFCWQRNFNAPSDRHPEDHYDTAILFTRCNF 337
QY 322 CG-ATTCDTLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDMNVKCEEV 380
Db 338 CGKEQCDTLGMADVGMCDPKRSCSVIKDEGLQAQAVTLAHELGHVLSMHPDSDKPCVRL 397
QY 381 FGKURAHMNSPTLIQIDRANPWSACSAAITDPLDSHGDCDLDOPSKEPISLPEDLFGA 440
Db 398 FGPNGKYHMAPEFIHVNKTLPMSPCSAVYLTLLDDHGDCLLDAPTSLVPLPTGLFGH 457
QY 441 S--YTLQOCELAFLAVGSKPCP---YMOYCTKLMCTGKAKQMVQCTRH--FFWADGTSC 493
Db 458 STLYELDQOQCKQIFGPDFRHCNPTSVEDICVOLCARHRSDEPICHKNGSLWADGIPC 517
QY 494 GEGKCLKAGCVERHNLN--KHRVDGSAWAKWDYGPCSRTCGGGVQLARRQCTNTPANG 551
Db 518 GPHGLCLDGSCLVKEDVENPKAVVVDGWPWRPWGQCSRTCGGGIOFSNRECDNMPGNC 577
QY 552 GKYCEGVVRYKSCNLEPCPSSASGSKSFRQCEAFNGYNHSTNRLTLAVAVPVKYSGV 611
Db 578 GRFCLGERVKYQSCNTEECPP--PNGKSFRCQCEKYAYNH-TDLNGNFWLOWPKYSGV 634
QY 612 PRDKCKLICRANGTYFYVLAPKVVVDGTLCSPDSTSVQVQKCIKAGCDNGLSKRFRD 671
Db 635 PRDRCKLCFARGRSEFKVPEAKVIDTGLCPDFTLSICVRGQCVKAGCDHWVNSPKKLDK 694
QY 672 CGVCGGDNKSKCKVTGLTFRPMHGNFVVAIPAGASSIDIRQYKGLIGDNNYALKNS 731
Db 695 CGVCGGKGTACRKTSGSFTPFPSYGNIDIVIPAGATNIDVKQRSHPGVRNDYLAUKTA 754
QY 732 QGKYLNGHVFVVSVERDLVVKGLLRYSGTGTAVESLQASRPILPTEVLSV-GKMT 790
Db 755 NGQVLLNGNLAIQEDILVKGLIYSSMATLERLQSFQALPELTVQLTVSGEVF 814
QY 791 PPRVRYFYLKPEPREDKSSHHPKDRGSPVLHNSVLSNQVEQDDRRPARWAGSWG 850
Db 815 PPKVRYTFFVEND--MDFSVQNSKERATT---NIQSL-----PSAEWVLGDSWE 859
QY 851 CSASCGSLQKRAYDRCGSAQRTVPACDAHARPVETOACG-EPCP 895
Db 860 CPSTCRGSQWRRTVECDPSPQAS-DTCDEALKPEDAKPCGSPCP 904
```


QY 361 HELG 364
Db 361 HELG 364

RESULT 15

US-10-174-590-352
; Sequence 352, Application US/10174590
; Publication No. US20030008352A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430RIC42

; CURRENT APPLICATION NUMBER: US/10/174,590

; CURRENT FILING DATE: 2002-06-18

; Prior application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 352

; LENGTH: 837

; TYPE: PRT

; ORGANISM: Homo Sapien

; ORGANISM: Homo Sapien

US-10-174-590-352

Query Match 37.1%; Score 1914; DB 9; Length 837;
Best Local Similarity 46.3%; Pred. No. 2.7e-128;
Matches 381; Conservative 123; Mismatches 267; Indels 52; Gaps 16;

QY 1 MLLGLILTAAGTAGGSEPEREVVPIRLDPDINGRIRYWRGPDSDGGLIFQITAF 60
Db 37 LLLLLLASLLPSARLASPLPREEEIVFEKLGNSVL-----PGSGAFARLLCKLQAF 88
QY 61 OEDFYHLHLPDAQFLAPAFSTHGLVPLQGLTGGSSDLRRCFYSGDVNAEFDSPFAVSLC 120
Db 89 GETLLELEQDSGVQVEGLTVQLGAPE-LLGGAEP--GTYLTGTINCEPESVASLHWD 145
QY 121 GG-LRGAFYRGAEYVISPLPNASAPAAQNSOGAHLLORRGVPGPGSGCTHTSRGCVASG 179
Db 146 GGALLGVLYRGAEHLQLPLEGTPNSA--GGPGAHILRRK-----SPASSGGPWCNV--- 196
QY 180 WNPAILRALDPYKPRRAGGESRRRRSGRAKRFVSIPIRYVETLVVADPSMVKFRGADIE 239
Db 197 -----KAPLGSPSPRPR--RAKRFASLSRFVETLVVADDKMAAFHGAGLK 239
QY 240 HYLLTLATAARYRHPSTILNINIVVKKLLRRDSDGPKVTGNAALTLRNFCAMQKLL 299
Db 240 RYLLTYMAAAAKAFKPSIRNPVSLVTVLILGSGEEGPQVPSAAOTLRSFCAMQKLL 299
QY 300 NKVSDKHPEYWDPAILFTQDLGGATTCTLGMADVMTCDPKRSCSVIEDDGLPSAFTT 359
Db 300 NTPEDSGPHFDFTAILFTQDLGGATTCTLGMADVMTCDPKRSCSVIEDDGLPSAFTT 359
QY 360 AHELGHVFNHNDNVKVEEVEFKL-RANHMSPTLIQIDRANPWSACSAAIITDFLDSG 418
Db 360 AHELGHVFNHNDNSKPCISLNGPLSTSRHVMAPVMAHVDPEEPWSPCSAREFITDFLNG 419
QY 419 HGDCILLDQSPKPSLPEDLPGLASYTLSSQCELAFGVSKPCPYM-QYCTKLMCTCKAKGQ 477
Db 420 YGCHLLDKPEAPLHLVPTTPGKYDADQCQLTFGPDSDRHCPCPLPPCAALWCSGHLNGH 479
QY 478 MVQOTRHFVWADTSGEGKCLCKGACVERHNLNKHVD--GSAKWDPDYPGCSRTCGGG 535
Db 480 AMCQTKHSWADTGPCPAQACMGGRCLHWDQLQDENIFQAGGWWGPGWPGDCSRTCGGG 539

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 18:46:45 ; Search time 3590 Seconds
(without alignments)
12870.675 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853
Sequence: 1 atgctctgctgagcatcct.....gcgtctgagcgctgctga 2853

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estchm:*
3: em_estcin:*
4: em_estimu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query Match	Length	DB ID	Description
1	632.4	22.2	2931	11 BC003269	BC003269 Mus muscu
2	530.2	18.6	849	13 B1103177	B1103177 602889417
3	511	17.9	681	13 BG921734	BG921734 602825569
4	409.2	14.3	961	10 BB612189	BB612189 BB612189
5	391.4	13.7	522	10 BE553572	BE553572 ut47B04.Y
6	345.6	12.3	544	12 BF078689	BF078689 229090 MA

7	339.4	11.9	627	10 BE281680	BE281680 601099657
8	270.6	9.5	654	10 BE016461	BE016461 F733410.X
9	263.4	9.2	738	13 B1819909	B1819909 603035454
10	254.8	8.9	746	13 B1733795	B1733795 603351940
11	245.8	8.6	661	10 BB220373	BB220373 BB220373
12	244.2	8.6	660	10 BB651049	BB651049 BB651049
13	242	8.5	662	10 BB642806	BB642806 BB642806
14	240.2	8.4	951	14 B0930852	B0930852
15	229.8	8.1	945	12 BF033017	BF033017
16	227.8	8.0	945	12 BF033017	BF033017
17	226	7.9	984	17 BMS70576	BMS70576
18	224.4	7.9	797	14 BQ58765	BQ58765
19	222.6	7.8	305	12 BF320986	BF320986
20	219.6	7.7	598	13 BM069515	BM069515
21	215	7.5	674	10 BM225365	BM225365
22	209.4	7.3	723	12 BG826611	BG826611
23	209.2	7.3	949	14 BQ893675	BQ893675
24	206.2	7.2	652	9 A1400905	A1400905
25	206	7.2	420	10 AW353152	AW353152
26	205	7.2	681	12 BE912007	BE912007
27	202.8	7.1	562	13 BJ009737	BJ009737
28	201.2	7.1	602	10 AM137467	AM137467
29	198.6	7.0	601	14 BQ332355	BQ332355
30	196.4	6.9	680	12 BF446232	BF446232
31	195.6	6.9	528	9 AA288689	AA288689
32	193.2	6.8	518	17 AZ374992	AZ374992
33	188	6.6	770	17 CANS03W7R	CANS03W7R
34	186.4	6.5	689	12 BG242316	BG242316
35	185.6	6.5	667	10 BB624639	BB624639
36	185.4	6.5	522	10 AW046509	AW046509
37	184.8	6.5	640	12 BE742606	BE742606
38	183.8	6.4	504	12 BG553803	BG553803
39	181.6	6.4	543	14 B0011430	B0011430
40	180.6	6.3	468	10 BE646878	BE646878
41	179	6.3	475	9 A1148739	A1148739
42	178.4	6.3	264	10 BB576723	BB576723
43	177.4	6.2	566	10 AV617111	AV617111
44	174.6	6.1	257	9 A1644268	A1644268
45	172.4	6.0	464	10 BE666088	BE666088

ALIGNMENTS

RESULT 1
BC003269
LOCUS
DEFINITION
Mus musculus, similar to a disintegrin-like and metalloprotease (repolysin type) with thrombospondin type 1 motif, 1, clone IMAGE:3499206, mRNA.
ACCESSION
BC003269
VERSION
BC003269.1 GI:14707747
KEYWORDS
HTC.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2931)
Srausberg, R.
Direct Submission
Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA.

NTF-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-MGSC

REMARK

COMMENT

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amebcm.tmc.edu
 Gunatane, F.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILIN at: <http://image.llnl.gov>
 Series: IRK Plate: 9 Row: b Column: 7
 This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

1. 2931
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="IMAGE:349206"
 /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI_CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector::pCMV-SPORT6"

BASE COUNT 708 a 812 c 811 g 600 t
 ORIGIN

Query Match 22.2%; Score 632.4; DB 11; Length 2931;
 Best Local Similarity 61.0%; Pred. No. 1.5e-125;
 Matches 1083; Conservative 0; Mismatches 681; Indels 12; Gaps 3;

631 AAGCGTTGCTGCTATCCCGGCTAGCTGAGAGCGTGTGTGCGGCGAGTCAATG 630
 1144 AAGGATTTGTGTCACCCCGCTTATGTGMAACCATGCTGCTGACGACGATCATG 1203
 691 GTCAAGTTCCAGCGCGGAGCGAGTCAATTCCTGCTGACGCGGCGGCGAGCGGCG 750
 1204 GCCACTTCCAGCGGACGAGCTTAACCATTTACCTTTACCTGTTCTGCTGCGAGCC 1263
 751 CGACTACCGGCTCCAGCATCTCAACCCATCAACATCTGTTGTGTCAGAGTGTG 810
 1264 AGGTTTAAAGCATCCAGATAGAGATTCATTCCTGCTGCTGCTGCTGCTGCTGCT 1323
 811 CTCTTATGAGATGCTGATCGGCGGCGGAGTCAACGCGGCAATGCGGCTGCTGCTG 870
 1324 GTCAATATGAG 1383
 871 AACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
 1384 AATTCTGCACTGCGAGAAACACACACAGCCCGAGTACCGGAGTCCAGAGCACTAT 1443
 931 GACACTGCACTCTCTTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 990
 1444 GACACTGCAATTTCTGTCACAGAGAGATTTATGTGCTCCACAGTGTGACACTCTC 1503
 991 GGCATGGCTATGTTGGTACCATGTGACCCCAAGAGAGAGAGAGAGAGAGAGAGAG 1050
 1504 GGGATGGAGATGTTGGAACCTGATGTGACCCAGAGAGAGAGAGAGAGAGAGAGAG 1563
 1051 GATGGCTTCATCAGCTTACCACTGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1110
 1564 GATGGTTTGCAGAGCGCTTACACAGCCAGCAATTTGGGCCATGTTTAAACATCCG 1623
 1111 CATGACATGGAAGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170
 1624 CAGCATGATGCTAAGCACTGTGCGAGTGTGATGTGATGTGATGTGATGTGATGTG 1683
 1171 TCCTCCAGCCCTCATCAGATGAGCCGCTGCAACCCCTGTCAGAGCTGCTGCTGCTG 1230
 1684 GCTTCGATGCTCTCAGCTTACACATAGCAGCCCTGCTGCTGCTGCTGCTGCTGCTG 1743
 1231 ATCAACGACTTCTGAG 1290
 1744 GTCAAGTCTTCTGATATAGACACGAGGAGATTTGATGAGCAAGCCCAAGAAATCA 1803
 1291 ATTCCTGCTGCGGAGAGATCTGCGGCGGCGGAGTCAACCTGAGCCAGCACTGTGAG 1350

1804 ATCAAGTCTCTTCTGATCTTCCCGGTACCTGTGATGATGCAACCGGAGTCACTT 1863
 1351 GCTTTGGCGGTGGCTCCAGAGCCCTGTCTTACATGACATGAC---TGACCAACCTGTG 1407
 1864 ACATTGGAGAGAGATCCAGACACTGCTGATGACAGCCACACATGATGATGATGATG 1923
 1408 TCCACCGGAG 1467
 1924 TCCATGTCGACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1983
 1468 GGCACCACTGTGTGAG 1527
 1984 GGCACCACTGTGTGAG 2043
 1528 CTCACAGACACAG-----GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1581
 2044 ATGAAGCATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103
 1582 TGCCTGCGACATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1641
 2104 TGCCTGCGACCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2163
 1642 CCTGCAACGAGGAG 1701
 2164 CCAAG 2223
 1702 GAGCCCTGCGGAG 1761
 2224 GAG 2280
 1762 AAGGCTACACAG 1821
 2281 AATGATTTTCCAAAGCTTCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
 1822 TCCGCGCTGTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1881
 2341 GCGGCGCTGTCTCCAAAG 2400
 1882 TTTCTATGCTGCTGAG 1941
 2401 TTTTTCGCTTACAGCCAG 2460
 1942 GTCTGTGTCAG 2001
 2461 GTCTGTGTCAG 2520
 2002 AGATTGCAAG 2061
 2521 AGATTGCAAG 2580
 2062 CTCTTACCAAG 2121
 2581 ATAGTCACTAGTACAG 2640
 2122 AGCATGATATCCGAG 2181
 2641 AACATTGAAGTAAACATGGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
 2182 CTGAAG 2241
 2701 ATTAGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
 2242 CGGAG 2301
 2761 CAAG 2820
 2302 AGCTTGAAG 2361
 2821 AGAATTCGAG 2880
 2362 AAGATGACAG 2397

Db 2881 CATGCTCTCCGACCAAAATTAATTCACCTACTT 2916

RESULT 2
BI103177 849 bp mRNA linear EST 26-JUN-2001
LOCUS 602889417F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044493
DEFINITION 5', mRNA sequence.
ACCESSION BI103177
VERSION BI103177.1 GI:14554070
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 849)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: L14M1122 row: a column: 06
High quality sequence stop: 743.
Location/Qualifiers
1..849
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="5044493"
/lab_host="NCI_CGAP_Kid14"
/note="Organ: kidney; Vector: pCMV-Sport6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library. 1"
BASE COUNT 193 a 250 c 250 g 156 t
ORIGIN

Query Match 18.6% Score 530.2; DB 13; length 849;
Best Local Similarity 85.0% Pred. No. 1e-103;
Matches 698; conservative 0; Mismatches 113; Indels 10; Gaps 9;

QY 1189 ATCGACCGTGGCAACCC-TGGTCAGCTGCGAGTGCATCATCAC-CGACTTCTCGG 1246
Db 26 ATAGACCGGCGCAACCCCTGGTGTGAGCTGAGCGCTGCATTATTAACGTCCTCGG 85
QY 1247 ACAGGGGGGAGGTGACGCTGCTGGAGCAACCCAGCAAGCCCTCCGCGGAGG 1306
Db 86 ATAGTGGGATGTGATGCTGCTGTGGAGCAACCCAGCAAGCCCTGCTGAGG 145
QY 1307 ATCTGCGGGGCGCAGCTACACCTGAGCAGCACTGAGCTGCTTTGGCGTGGCT 1366
Db 146 ACCGCGGGGCGACAGCTACATTGTGAGCAGCACTGAGCTTTGGGGTGGCT 205
QY 1367 CCAAGCCTGTCTTACATGACGATGACCAAGCTGTGTGACCGGGGAAGGCCAAG 1426
Db 206 CTAACCCCTGCCATATATGCTACTGTACAAACCTGTGTGCTGCTGCAAGGCCAAG 265
QY 1427 GACAGATGTTGTGAGAGCCGCGCCTTCCCTGGCGAGTGGCAAGCTGTGGCGAG 1486
Db 266 GGCATATGTGTGCGAGACTGCGCACTTCCCTGGCGAGTGGCGAGCTGTGGGGAAG 325
QY 1487 GCAAGCTGTGCTCAAAAGGGGCTGCTGAGAGACACAACTTCAACAGCACAGAGGTG 1546
Db 326 GCAAGTGTGCTCAAAAGGAGGCTGCTGAGAGACACAACTTCAACAGTACCGGGTGG 385

QY 1547 ATGTTCTCGGCGCAAAATGGATGCTATGAGCCCTGTGCGCAGATGTGTGGGCG 1606
Db 386 ACGGCTCTTGGGCGCAAGTGGGAGCCCTTAGCTTCTGCTCCGCACTGCGGTGGGGCG 445
QY 1607 TCGACCTGGCGAGAGGAGCTGACCAACCCACCC-CGCCAAGGGGGGCAAGTACTGC 1665
Db 446 TCGACCTGGCGGAGGAGCTGACCAACCTTACCCATGCAACGGTGGAAATACTGC 505
QY 1666 GAGGAGTGAAGGTGAATACCAGTCCGCAATGTGAGACCTCCCGCAGCTAGCTCC 1725
Db 506 GAGGAGTGAAGTGAATACCAGTCCGCAATGTGAGACCTCCCGCAGCTAGCTCC 565
QY 1726 GAAAAGAGCTTCGGGAGAGCAGTGTAGGCTTTCAACGGCTCAACACAGCAACAC 1785
Db 566 GGAAGAGCTTCGGGAGAGCAGTGTAGGCTTTCAACGGCTCAACACAGCAACAC 625
QY 1786 CGGCTCACTTCGCGGTGGCATGTGGTCCCAAGTACTCCGCGCTGTCTCCCGGAGCA 1845
Db 626 CGGCTCACTTACGTTGTGTGATGTGGTACCCAAAGTACTCAGGCGGTGTACACGTGAC 685
QY 1846 TCGAAGCTCATCTGCCAGCAATGAGCAGTGTGCTATGTGCT-GGCACCCAAAGT 1904
Db 686 TGTAGGCTCATCTGCCAGC-ATGGAGCTGGCTACTTATGTACTAGACACCTAAGGT 744
QY 1905 GGTGGA-CGGCAGCCTGTG-CTCTCGTACCTCCAGCTCCGCTGTGTGCAAGGCAAG-TG 1961
Db 745 GGTGAGCCGATGAGCTGTGTACACCTGACCTCCGCTGTGTGTGCAAGGCAAGT 804
QY 1962 CATCAA-GGCTGCTGTGATGAGCACTGGGCTCAAGA 2000
Db 805 CATCCAGCGGCTGTGCGAAGCGCAAACTGGGCTCAAGA 845

RESULT 3
BG921734 681 bp mRNA linear EST 05-JUN-2001
LOCUS 602825569F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954456 5',
DEFINITION mRNA sequence.
ACCESSION BG921734
VERSION BG921734.1 GI:14302210
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 681)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: L14M0915 row: i column: 17
High quality sequence stop: 680.
Location/Qualifiers
1..681
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4954456"
/lab_host="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/note="Organ: mammary; Vector: pCMV-Sport6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.


```

/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:3153391"
/clone_11b="NCL CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by life technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

0

COMMENT Contact: Smith RPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithr@mail.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
REVERSE: GTTTCACAGCAGCAGC
Plate: 43 row: P column: 17
Seq primer: ATTAGTGTACATATAG.
Location/Qualifiers
1. 544
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 117 a 154 c 169 g 104 t

ORIGIN

Query Match 12.3%; Score 349.6; D3 12; Length 544;
Best Local Similarity 88.6%; Pred. No. 6.4e-65;
Matches 379; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 559 GCCCTGACCTTACAAAGCGCGGCGGCTTCGGGAGAGTCGATGCGCGCGCAGG 618
Db 544 GCCCTGACCTTACAAAGCGCTCCGGATAGGCTTAGGGGAGTCGACGCCGCGCAGG 485

QY 619 TCTGGCGCGCCAAAGCGTTCTGCTATCCGCGGTACGTGAGAACGCTGTGTCGG 678
Db 484 TCCGGCGCGCCAAAGCGCTTCGTATCCGCGATAGTGAACGCTGTGTCGCA 425

QY 679 GAGAGTCATGTGTCAGTTCATCCAGCGCGGAGCCTTGACATTTACTGTGACGCTGCTG 738
Db 424 GAGAGTCATGTGTCAGTTCATCCAGCGCGGAGCCTTGACATTTACTGTGACGCTGCTG 365

QY 739 GCAACGGCGCGGAGCTTACCGCCATCCGACATCTTCAACCCATCAACATCGTTGTG 798
Db 364 GCAACGGCGCGGAGCTTACCGCCATCCGACATCTTCAACCCATCAACATCGTTGTG 305

QY 799 GTCAAGTGTGCTCTTATAGATCGTGAATCCGCGGCGCAAGTACCGGCAATCGCGC 858
Db 304 GTCAAGTGTGCTCTTATAGATCGTGAATCCGCGGCGCAAGTACCGGCAATCGCGC 245

QY 859 CTGACGCTGCGCAACTTCTGTGCTGCGCAAGAAAGACTGACAAATGATGACAAAGCAG 918
Db 244 TTGACGCTGCGCAACTTCTGTGCTGCGCAAGAAAGACTGACAAATGATGACAAAGCAG 185

QY 919 CCGAGTACTGGGACACTGACATCTCTTACACAGGAGAGACTGTGGAGCCCAAGC 978
Db 184 CCGAGTACTGGGACACTGACATCTCTTACACAGGAGAGACTGTGGAGCCCAAGC 125

QY 979 TGTGACAC 986
Db 124 TTGAGACCC 117

RESULT 7
BE281680 627 bp mRNA EST 26-OCT-2000
LOCUS BE281680 601099657f1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3491991 5',
DEFINITION mRNA sequence.
ACCESSION BE281680
VERSION BE281680.1 GI:9156727
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 627)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM8537 row: a column: 16
High quality sequence stop: 530.
Location/Qualifiers
1. 627
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="IMAGE:3491991"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary."
/lab_host="DH10B"
/note="Organ: Lung; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 153 a 159 c 177 g 138 t

ORIGIN

Query Match 11.9%; Score 339.4; DB 10; Length 627;
Best Local Similarity 83.6%; Pred. No. 1e-62;
Matches 504; Conservative 0; Mismatches 86; Indels 13; Gaps 10;

QY 1850 AGCTCATTCGCGAGGACATGACGCTGCTATATGTCGCGACCCCAAGGAGTGGT 1908
Db 1 AGCTCATTCGCGAGGACATGACGCTGCTATATGTCGCGACCCCAAGGAGTGGTGG 60

QY 1909 GACGCGACGCTGCTCTCTCTGAC--TCCACTTCGCTCTGTGTCAGAGCAAGTGCATCAA 1967
Db 61 GAGTGTACCTGTGTACTTCTTCAATCCACCTCGCTGTGTGTCAGAGCAAGTGCATCAA 120

QY 1968 -GGCTGCTGTGTGAGGAACTGG--GCTCCAGAGAGAGATTGACAA--GTGTGGGCTGTG 2024
Db 121 TGGCTGCTGTGTGAGGAACTGGGAACTGTGTGTCAGAGAAATTTGACAAATGTGTGTG 180

QY 2025 TGGGAGAGACAAATA--GAGCTCAAGAGAGTGTGACTCTTCACCAAGCCATGCATG 2083
Db 181 TGGTGTGAGACAAATAAGACCTTAAGAGGTGACAGGACTCTTCACCAAGCCATGCACG 240

QY 2084 GCTACAAATTCGTTGGGAGGACATCCCGAGGAGGCTCAAGACTGACATCCCGAGGCGG 2143
Db 241 GCTACAAATTTTGTAGTGGACATCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

QY 2144 GTTACAAAGGCTGATGCGGGATGACAACTACTGCTGTGAAGAACGCAAGCAAGT 2203
Db 301 GTTACAAAGGCTGATGCGGGATGACAACTACTGCTGTGAAGAACGCAAGCAAGT 360

QY 2204 ACCGTGCTCAAGGCAATTCGTTGGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2261
Db 361 ACCGTGCTCAATGAGGCACTTAGTGTATCCGCTGTAGAGGAGGAGGAGGAGGAGGAGGAGG 420

QY 2262 CAGCTGCTGCGGTACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2321
Db 421 CAGTGTGTACGATATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 479

QY 2322 CATCTGTGAGCGGCTGACCGGTGAGGCTCTCTCCGTGGGGAAGATGACACCGCCCGGCT 2381
Db 2322 CATCTGTGAGCGGCTGACCGGTGAGGCTCTCTCCGTGGGGAAGATGACACCGCCCGGCT 2381

Df	480	CATCCTTGAGGACACTACCGGTGGAGGTCCTGACGGTGGGGAAGTAGCACCCAGACC---	GT	536
QY	2382	CCGCTACTCCTTTATCTGCCCAAGAAGACCTCGGAGAGACAAGTCCTCATCCTCCAGGA	2441	
Df	537	GTCGGTATTCTACTACCTGCACAAAGAACCTTC-	GGAGGACAAAGTCATCGCCCCCAAGAC	595
QY	2442	CCC	2444	
Df	596	CTC	598	
RESULT 8				
LOCUS	BE016461			
DEFINITION	f473d10.x1 zebrafish Research Genetics C32 fin Danio rerio cDNA 3' similar to TR:054768 054768 A DISINTEGRIN AND METALLOPROTEASE WITH THROMBOSPONDIN MOTIFS ; mRNA sequence.			
ACCESSION	BE016461			
VERSION	BE016461.1			
KEYWORDS	GI:8280907			
SOURCE	EST.			
ORGANISM	zebrafish.			
	Danio rerio			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes			
	; Cyprinidae; Danio.			
	1 (bases 1 to 654)			
REFERENCE	Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy			
AUTHORS	S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood			
	,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,			
	Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,			
	Kohn,S., Shih,T., Jackson,Y., Cardenas,W., McCann,R., Waterston,R.			
	and Wilson,R.			
TITLE	Mashu Zebrafish EST Project 1998			
JOURNAL	Unpublished (1998)			
COMMENT	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbraflsh@watson.wustl.edu CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Research Genetics. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Research Genetics web address: http://www.researchgenetics.com/ Seq primer: 17 Et from Amersham High quality sequence stop: 489. Location/Qualifiers 1..654 /organism="Danio rerio" /db_xref="taxon:7955" /clone_lib="Zebrafish Research Genetics C32 fin" /tissue_type="Fin" /lab_host="GenEHogs (HS996, a phage-resistant isolate of DH10B)" note="Vector: pUT7T3D-Pac with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was prepared from zbraflsh(C32) fin, and was then primed with a Not I -oligo(drf) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library is non-normalized. Library was constructed by Ning Wu. NOTE:C clones from this library are only available thru Research Genetics (www.resgen.com)."			
FEATURES	source			
BASE COUNT	121 a	193 c	205 g	135 t
ORIGIN				
Query Match		9.5%;	Score 270.6;	DB 10; Length 654;
Best Local Similarity		63.4%;	Pred. No. 6,7e+8;	
Matches 414;	Conservative	0;	Mismatches 239;	Indels 0; Caps 0
QY	853	GGCGCCCTAGACGCTGGCACAATTCTGCTCGGAGAGAAGCGAGACAAAGTGAAGAC	912	
Df	1	GGCGCCCGCGACGCTTCAGACACTTCTTGGAAGTGGGAGAACCAAGACGACACCTCTGAGATGAC	60	

OY	913	AAGCAACCCCGAGTACTGTGGACACTGCCATCCTCTTCACCGAGGACGCCTGTGGAGGCC	972
Dd	61	GATCATCAGATATCACATCCATGACGGCGGATCCCTTTACCACGGACGGATTCTGTGGGGCAT	120
OY	973	ACGACCTGTGACACCCCTGGGCGATGGCTGATGTGGGTGCCAATGTGTGAACCCCAAGAGAAGC	1032
Dd	121	CACCTGTGTGACACTCTTGCGGATGTGGCGATGTAGAACTGTGTTCGCCAGAGAGGACG	180
OY	1033	TGCTCTGTCAATTTGAGGACGATGGGCTTCATCAGCCTTTACCACTGTGCCACGACCTATGGGC	1092
Dd	181	TGTGCTATCATTTGAGGACGAGCGGCTCGACGCTCCTTCACTGATGTGGCTCATGAGATGGGT	240
OY	1093	CACGTGTTCAACATGCCCCCATGACAATGTGAAGCTGTGTGAGAGGTGTGGGAACCTC	1152
Dd	241	CACCTGTGGGTCTGTCTCCACGACGACTCCAAAGTTCTGTAGAGAGCGTTTGGCTCCAGC	300
OY	1153	CGAGCCAACCAACATGATGTGCCCGGACCCCTCATCCAGATCGACCGCTCCACACCCTGGTCA	1212
Dd	301	GAGGACCAAGCGGGTGTATGTCTCTCATCTGCATCTCATCGACGCGCTCCAAACCCCTGGAGC	360
OY	1213	GCCTGTCACTGCTGCCATATATCAGCAGACTTCTGTGACACAGGGGGCACGGTACTGGCTGCTG	1272
Dd	361	CGCTGTACACCTCCACACCATCATCCGACTCTTTCGACGACGAGAAACGGGAGTGTCTCTG	420
OY	1273	GACCAACCCAGAACCCCATCTCCCTGTGCCGCGAGATTCGCCGGGCCACAGTCAACCCCTG	1332
Dd	421	GACGCTCCCCTGGTTCCTCTGTCTGGGGCGGAGGAGCTGCCGGGTCAAGATTAAGATGCG	480
OY	1333	AGCCAGACGATGCGAGCTGGCTTTTGGCGGTGCCCTTCACACCCCTGTCTTACATGTAGATAC	1392
Dd	481	GTGACGACGATGGCGCTGTGGCGCTTTCGGCTCAGAGTACATGTGTGTGCCGGGTGAGATTC	540
OY	1393	TGCACCAACGCTGTGGTGGACCGGGAAGGCCAAGAGACATGTGTGCCAGACCCGCCAC	1452
Dd	541	TGCGCTGTGGCTGTGGTGTCTCTGTCTCATCAGAAAGGGTCAATGTGTGTGTGACCAAGAAG	600
OY	1453	TTCCCTGTGGGCGGATGGACACGACTGTGTGGCGAGGCGCAAGCTTGCCCTCAAAGC	1505
Dd	601	GTGGCCGCTGTGATGAGACGCCATGTGTGGAGGCGCAGGATCTGTCTCGAGGG	653
RESULT 9	BtB19909	738 bp mRNA	EST 04-OCT-2001
LOCUS	60303545AF1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176476 5'		
DEFINITION	mRNA sequence.		
ACCESSION	BtB19909		
VERSION	BtB19909.1 GI:15931459		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 738)		
JOURNAL	NIH-MGC http://mgl.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLML1439 row: 1 column: 13 High quality sequence stop: 735. Location/Qualifiers 1. 738 /organism="Homo sapiens" /db_xref="taxon:9606"		

```

/clone="IMAGE:5176476"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pcmv-sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age 69. Library is
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dt primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC library."

```

BASE COUNT 169 a 204 c 225 g 140 t

ORIGIN

Query Match 9.2%; Score 263.4; DB 13; Length 738;

Best Local Similarity 63.1%; Pred. No. 2.4e-46;

Matches 440; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

```

QY 1715 GCTCAGCCTCCGGAAGAGTTCCGGAGAGCAGTGTAGCTTCAACGGCTCAACC 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 GCGCCCTGAGGGAAGAGCTTCAGGAGCAGAGTGTAGATATATAGCTTCAATT 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1775 ACAGCACCACCGCTCCTCTCCGCGTGCATGGTGGCCCAAGTACTCCGCGTCTC 1834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 ACAGTACATGACGGGAATCT---CCTGAGTGGGTCCCAAGTATGCTGGGTCTCC 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1835 CCGGAGCAGTGCAGCTCATCTGCGGAGCCCATGGCTACTTCTATGTGCTGG 1894
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 CCGGAGCAGCTCATCTGCTGCGGAGCCCGGAGGAGCAGTCAAGAGTTCG 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1895 CACCCAGGTGTGAGCAGCAGCTGCTCTCTCTACCTCCACCTCCGCTGTCCAG 1954
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 AGGCGCAGTGTATGATGACACCTGTGTGGGCAAGAACACTGGCAGTCTGTCCGTG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1955 GCAGTGCATCAAGCTGGCTGTGATGGAGCACTGGGCTCCAGAGAGATTCGCAAGT 2014
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 GCGAGTGTGTAAGCGCGGCTGTGACCATGTGTGCTGCTCGGAGCTGGACAAAT 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2015 GTGGGCTGTGTGGGAGACATTAAGAGTGCAGAGAGTGTGAGTCTTCAACAGC 2074
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 GCGGCTGTGTGGGAGAGCACTCCTGCGAGAGATGTCGCGGTCTCCACCCCA 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2075 CCATGATGCTCATATTTGTGTGGCCATCCCGGAGCCCTCAAGCATGACATCC 2134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 CCATTTATGCTCAATGACATTTGTCACCATCCAGCTGTGCTCAATATTTGACGTGA 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2135 GCCACGCGGTTCAGAAAGGCTGATGGGAGTACAACTACTGCTCTGAGAGACAGCC 2194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 AGCAGCGGAGCCACCGGGGTGTGCAAGAGATGGAGTACACTGGCGCTGAGAGAGCTG 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2195 AAGGCAAGTACCTGCTCAACGGGCAATTTGTGTGCTGGGCTGAGAGCGGACCTGGTGG 2254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 ATGGGCGTGTACGCTCAACGGGCAACCTGCTCTCTCAATAGAGGACATCTTGG 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2255 TGAAGGCACTGCTGCTGCTGAGAGGAGCAGGAGCAGAGTGTGAGAGAGCTGAGAGCTT 2314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 TGAAGGAGGACATCTGAGAGTACAGGCGCTCCATGAGACACCTGGAGCGCTGAGAGCT 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2315 CCGGAGCCATCTGAGAGCGCTACCGGTGAGAGTCTCTCTCCCT---GGGAGAGTACAC 2371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 TCGGCGCTTGGCAGAGCCCTGACAGTACAGTCTCTGACAGTCTCCCTGCGGAGGCTTCC 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2372 CGGCGCGGCTCGCTACTCTTCTATCTGAGCCAAAGA 2408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 CCGCAAAAGTCAATACACCTTCTGTGTACTATATGA 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10
LOCUS B1733795 746 bp mRNA linear EST 20-SEP-2001
DEFINITION 603351940F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5359712 5',

```

ACCESSION mRNA sequence.
VERSION B1733795
KEYWORDS B1733795.1 GI:15710808
SOURCE EST
ORGANISM house mouse.
MUS musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 746)
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
Plate: L1AM11914 row: 0 column: 09
High quality sequence stop: 746.
Location/Qualifiers
1..746
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5359712"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/notes="Organ: eye; Vector: pcmv-sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

```

FEATURES

source

BASE COUNT 213 a 160 c 216 g 156 t 1 others

ORIGIN

Query Match 8.9%; Score 254.8; DB 13; Length 746;

Best Local Similarity 61.3%; Pred. No. 1.7e-44;

Matches 447; Conservative 0; Mismatches 273; Indels 9; Gaps 2;

```

QY 1463 CCGATGGACACAGCTGTGGCGAGGCAAGCTGCTGCTCAAGGGGCTGCGTGGAGAGC 1522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 CAGATGGACACAGCTGTGGCGAGAGTGTGTAAGTGGCAAGTGCCTGAGACAGA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1523 ACAACCTCAACA-----GCAGAGGTGATGTTCTTGGGCCCAATGGATGCTATG 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 CAGACATGAAGCATTTGTACTCTGTCATGAGAACTGGGAGACATGGGAGCGTGGG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1577 GCGGCTGTCCGCGACATATGTTGGGGGCGGTGCACTGGCCAGAGGAGCAGCCAAAC 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GAGACTGCTCAAGAACCTGTGTGTGGGAGTTCATACACATGAGAGATGTACAAAC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1637 CCAACCCCTGCAAGGGGGCAAGTACTGCGAGAGTGAAGGTGAATACCATCTGCA 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CAGTCCCAAGAAAGGAGGAGTACTGTGAAGCAAGATCCGTACAGGCTCTGTA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1697 ATGTGAGCGCTGCGCCACACTCAGCTCCGGAAGAGCTTCCGAGAGAGAGTGTAGG 1756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ACATCGAGAGCTGTCCAGAC---AATTAAGGAAAACGTTCAAGAGAGAGTGTAGG 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1757 CTTTCAAGGCTPACACACACACACCGGCTCACTCTGCGCGTGTGATGGTGGCCCA 1816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 CCGACATAGTGTTCACAAAGCTCTTNGGAAATGAGCCACATGTATAGAGTACACCA 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1817 AGTACTCGGCGTGTCCCGGAGCAAGTGAAGTATCTGCCGAGCCATGGCACTG 1876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 AGTACGCGGCGTGTCCGCAAGGAGAGTGAACCTGAGAACCCAAAGGCAATTG 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1877 GCATCTTATGTGCTGGCAGCCAGAGGTGTGAGGAGCGGTGCTCTCTGATGCA 1936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	418	GCTACTTTTTCGCTTACAGCCCAAGGTTTGAATGGACCTCCCTGATGCCAGACTCTA	477
QY	1937	CCTCCGTCGTGTCCAAAGCAATGTCATCAAGGCTGCTGTGATGGAAACCTGGGCTCCA	1996
Db	478	CCCTGTCTGTGTGCCAAGGGCAGTGTGTGAAGACGTGGCTGTGATCGCTCATATGACTCA	537
QY	1997	AGAAAGATTTGACAAAGTGTGGGTGTGTGGGGGAGCAATAAGAGCTGCAGAAAGTGA	2056
Db	538	AAAAGAASTTGGATTAAGTGTGGCGTTTGTGGAGGAAACGTTCCACATCAAGAAAGATGT	597
QY	2057	CTGACCTTTTACCAAGCCCATGCATGGCTGCACAAATTTGTTGGTGGCCATTCGCCGAGGGG	2116
Db	598	CAGGAATGTCACTAGTACAAAGACCTGGGTATCATGACATGTTCCAAATTTCTGCTGGAG	657
QY	2117	CCTCAAGCATGACATCCGCGAGCGCGGTTACAAAGGGCTGATCGGGATGCACAATACC	2176
Db	658	CCACCACATGAAATGAAACATCGGATTCAAAGGGGGCTCCAGAAACAATGGACACTTTC	717
QY	2177	TGGCTCTGA	2185
Db	718	TGGCTATTTA	726

RESULT 11	
LOCUS	BB220373
DEFINITION	BB220373 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA A530059019 3' similar to AF140673 Mus. musculus putative secreted metalloprotease ADAMT5 (Adamts5) mRNA, mRNA sequence.
ACCESSION	BB220373
VERSION	BB220373.2
KEYWORDS	GI:16353495
SOURCE	EST.
ORGANISM	house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 661)
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.

TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	On Jun 30, 2000 this sequence version replaced gi:8885326.

On Jun 30, 2010 this sequence version replaced y1:8885326.
 Contact: Yoshitake Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitoh-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapped-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Wael, K., Fujiwara, S., Inoue, K., Togawa, Y., Ixaka, M., Ohara, E.,
 Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama,
 S., Kawai, Y., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
,K., Fukuda,S., Hara,A., Itoh,K., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

location/Qualifiers
1..661
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="A530059019"
/clone_lib="RIKEN full-length enriched, adult male aorta
and vein"
/sex="male"
/tissue_type="aorta and vein"
/dex_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI. cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAAGATCCACAGACTCTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATTCGTGGATTAAATTAATATCCCCCCCCCCC
3']. cDNA was cleaved with xhoI and BamHI. Vector: a
modified pluescript KS(+) after bulk excision from Lambda
FLC I."

	BASE COUNT	ORIGIN
Query Match	160 a	167 c 180 g 154 t
Best Local Similarity 60.9%:		
Matches 400; Conservative		
QY 861 GACGTCGCCGAACCTTGCTGTGCCTGCAGCAAGAAGTCACAAGAGTAGTAGACAGACC		920
DB 4 GACCTCAGAAACCTTTTGCATAATGAGAGCACCAACATACCAAGCTAAGGAGATGATCAGCA		63
QY 921 CGAGTACTGGAGACATGCGCATCTCTTCAACAGGACGGACTGTGTGGAGCCACCACTCG		980
DB 64 AGAGCACTACGATGACAGCCATCTCTTCAACCCGAGAGGATTTATGTGGGCATCATTCATG		123
QY 981 TGACACCTCGGATGGCGATGTGGGTACCATGTGAGACCCCAAGAGAGCTGCTCTGT		1040
DB 124 TGACACCTCGGATGGAGTAGGACAGCTTGGAGACATATTTCTCCGAGCGCAGCTGTGCAT		183
QY 1041 CATTTGAGAGAGATGGGCTTCATCAGCTTACCACTGCCCCAGAGCTGGGCACAGCTGTT		1100
DB 184 GATTGAAGATGATGGCTTCATCAGCTTCACTGTGGCTCATGAATAATGGGCATTTACT		243
QY 1101 CAACATGCCCATGACATGTAAGTCGTGTGGAGGGGTTTGGAGAGCTCCGAGCCAA		1160
DB 244 TGGGCTTTTCATGACGATTCCAAATTTCTGTAAGAGAACTTGGTACTRACAGAAACAA		303
QY 1161 CCAACATGATTTCCCGACCTTCATCCAGATGCAAGCGCTGCCAACCCCTGTGACGCTGCA		1220
DB 304 GCCTTTATGTCTTCAATCTCTTACAGACATGATGATCCAMGCCCTGTCCMAATGCA		363
QY 1221 TGCTGCATATACGCACTCTCTGCAGACAGCGGCAAGGTGATGTGCTCTGTGACCAACC		1280

QY	1318	GCACGCTACACCCCGACCCAGCGAGTGGAGACTGGCTTTGGTCGCGGGGCTTCACGACCCCTGG	1317	
Db	545	CAGACCTCACGATGCCACCGACGAGCTGGAACCTTGACATTTGGCGCTAGTACTGGGTGTGC	604	
QY	1378	CCCTACATCATGAGTACTGTGCACCAAGCTGTGGTGCACCGGGAAGCGAAGG	1426	
Db	605	CCCTGCATGTGATGTCTGTGGCGCGGCTGTGGTGTCTGTGGTGGCGCCAGG	653	
RESULT 13				
LOCUS	BB642806			
DEFINITION	BB642806 RIKEN full-length enriched, adult retina Mus musculus cDNA clone A930039D02 5', mRNA sequence.			
ACCESSION	BB642806			
VERSION	BB642806.1			
KEYWORDS	GI:15402028			
SOURCE	EST.			
ORGANISM	house mouse. Mus musculus.			
REFERENCE	Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus. 1 (bases 1 to 662)			
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiemoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyu,T., Muramatsu,M. and Hayashizaki,Y.			
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagui,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Alizawa,K., Fukuda,Y., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y. Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues. Location/Qualifiers 1. 662 /organism="Mus musculus" /db xref="taxon:10090" /clone="A930039D02" /clone_lib="RIKEN full-length enriched, adult retina" /tissue="retina" /dev_stage="adult"			


```
Db 136 GGAAGATACCTGCTGGGTGGAGAGCCAGTACCACTCAAGCACGAGAGATGCCCC 195
QY 1714 AGCTCAGCCTCCGGAAAGACCTTCCGGAGAGACAGTGTAGAGCTTTCAGCGCTACAAAC 1773
Db 196 -----CTGACGCGGAAAGCTTCAGGAGCAGCAGTGTAGAGAGTATATGCTACAAAT 249
QY 1774 CACAGCACCAACCGGCTCACTCTCCGCGGTGGGATGGGTGCCCAAGTACTCCGCGGTGTCT 1833
Db 250 TACACTGACATGGAGCGGGAATCT---CTGCACTGGGTCCCAAGTATGCTGGGGTGTCC 306
QY 1834 CCCCAGGACAAAGTCAAGCTCATCTGCGAGCCAAATGCACTGGCTACTTCTATGTGTG 1893
Db 307 CCCCAGGACCGCTGCAAGTTGTCTGCGAGCCCGGGAGGAGCAGTTCAAAGTGTTC 366
QY 1894 GCACCCCAAGTGTGTGACGCGACGCTGTGCTCTCTGACTCCACCTCCGCTGTGTCCAA 1953
Db 367 GAGGCCAAGTGTATGTGACACCTGTGTGGGCCAGAAACACTGGCCATCTGTGTCCGT 426
QY 1954 GGCAGTGCATCAAGGCTGTGTGATGGGAAACCTGGGCTCCAAAGAGATTCGACAAAG 2013
Db 427 GGCAGTGTGTCAAGGCCGCTGTGACCATGTGTGTGACTCGCTCGGAAGCTGGACAAA 486
QY 2014 TGTGGGTGTGTGGGGAGACAAATAAGAGCTGCAGAGGTGACTGGACTCTTACCCAAG 2073
Db 487 TCGGGGTGTGTGGGGCAAAAGGCAACTCCTGCAGAGAGGTCTCGGGTCCCTCACCCCC 546
QY 2074 CCCATGATGGCTACAAATTCGTGTGGCCATCCCGCAGCGCCCTCAACCATCGACATC 2133
Db 547 ACCAATTATGGCTACAAATGACATTTGCACCATCCAGCTGTGTGCCACTAATATGTAGCGTG 606
QY 2134 CGCCAGGCGGTTACAAAGGGGTGATCGGGGATGATGACAACTACTGTGCTCTGAAAGACAGC 2193
Db 607 AAGCAGCG -GAGCACCCGGGTGTGCAGAAAGATGGGAAGTACTACTGGGCGTGAAGAGGT 665
QY 2194 CAAGGCAAGTACTGCTCAACGG 2216
Db 666 GATGGCGACTACTGCTCAACGG 688
```

Search completed: May 15, 2003, 22:52:54
Job time : 3616 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 04:02:47 ; Search time 4150 Seconds
(without alignments)
886.794 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853
Sequence: 1 atgcttctgtctgtgcatcct.....gcgtcttgaggcctgtctga 2853

Scoring table: OLIGO_NNUC
Gapop 60.0 , Gapext 60.0

Searched: 810007 seqs, 644969091 residues

Word size : 24

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications -NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2853	100.0	2853	10	US-09-965-631-3
2	2853	100.0	2853	10	US-09-965-631-3
3	1491	52.3	2469	9	US-10-163-316-3
4	1491	52.3	2940	9	US-10-163-316-1
5	1091	38.2	1104	10	US-09-965-631-5
6	959	33.6	966	10	US-09-965-631-1

ALIGNMENTS

RESULT 1
US-09-965-631-3
; Sequence 3, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION: Carl Johan
; APPLICANT: Fridtjof, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: NO. US20020115842A1el Human Proteases and Polynucleotides Encodin

```

FILE REFERENCE: LEX-0241-USA
CURRENT APPLICATION NUMBER: US/09/965,631
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/236,689
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2853
ORGANISM: homo sapiens
US-09-965-631-3

Query Match      100.0%; Score 2853; DB 10; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1  ATGCTTCTGTGCGGAGCTATACCTGCTTTCGCGGCGAGACCGCTGAGGCTGTGAG 60
1  ATGCTTCTGTGCGGAGCTATACCTGCTTTCGCGGCGAGACCGCTGAGGCTGTGAG 60
61  CCAGAGCGGAGAGTGTGCTTCCATCCGACTGAGACCGGAGCATTAACGGCGCGCTAC 120
61  CCAGAGCGGAGAGTGTGCTTCCATCCGACTGAGACCGGAGCATTAACGGCGCGCTAC 120
121 TACTGCGGCGGTCGCCGAGACTCCGGGATCAGGAGCTCATTTTTCAGATCAGACATTT 180
121 TACTGCGGCGGTCGCCGAGACTCCGGGATCAGGAGCTCATTTTTCAGATCAGACATTT 180
181 CAGAGGAGCTTTTACCTACCTGACGCGCGAGTCTCACTTGTGCTCCGCTTCTCC 240
181 CAGAGGAGCTTTTACCTACCTGACGCGCGAGTCTCACTTGTGCTCCGCTTCTCC 240
241 ACTGAGACATCTGGGCTCCCTCCAGGGGCTCACGGGGGCTTTCAGACCTGCGACGC 300
241 ACTGAGACATCTGGGCTCCCTCCAGGGGCTCACGGGGGCTTTCAGACCTGCGACGC 300
301 TCGTTCTATTTCGGGAGCTGTAACGCCGAGCCTGCTGCTGTGTGAGCCTGTGCT 360
301 TCGTTCTATTTCGGGAGCTGTAACGCCGAGCCTGCTGCTGTGTGAGCCTGTGCT 360
361 GGGGGGCTCCGCGAGACCTTTTGCTACGAGGCGCGGAGTGTATTAAGCCGCTGCC 420
361 GGGGGGCTCCGCGAGACCTTTTGCTACGAGGCGCGGAGTGTATTAAGCCGCTGCC 420
421 AATGTAAGGCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
421 AATGTAAGGCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
481 GGTGTTCGCGGCGGCTTCCGAGAGACCCACCTCTGCTGCGGGGTGGCTCGGCT 540
481 GGTGTTCGCGGCGGCTTCCGAGAGACCCACCTCTGCTGCGGGGTGGCTCGGCT 540
541 AACCCCGCATCTCTACGGGCGCTTGACCCCTTCAAGACCGCGGCGGCGGCTTCCG 600
541 AACCCCGCATCTCTACGGGCGCTTGACCCCTTCAAGACCGCGGCGGCGGCTTCCG 600
601 AGTCGTAAGCGGCGAGGCTGTGGGCGCGGAGGCTTGTGTATCCCGGCTACGCTG 660
601 AGTCGTAAGCGGCGAGGCTGTGGGCGCGGAGGCTTGTGTATCCCGGCTACGCTG 660
661 GAGAGCGTGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
661 GAGAGCGTGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
721 TATCTGTGACGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
721 TATCTGTGACGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
781 CCCATCAACATCGTTGTGTGCAAGGTGCTGCTTATAGAGATGTGACTCCGGGCGCAAG 840
781 CCCATCAACATCGTTGTGTGCAAGGTGCTGCTTATAGAGATGTGACTCCGGGCGCAAG 840

```

QY	841	GTCCCGGGCAATGGGGCCCGAGCGTGGGGCACTTGTGGCTGGGCAAGAACAGTGCAGC	900
Db	841	GTCCCGGGCAATGGGGCCCGAGCGTGGGGCACTTGTGGCTGGGCAAGAACAGTGCAGC	900
QY	901	AAATGAGTGAACAAGCAACCCCGAGTACTGGGACACTGCATTCCTTTGACCAAGCGAGAC	960
Db	901	AAATGAGTGAACAAGCAACCCCGAGTACTGGGACACTGCATTCCTTTGACCAAGCGAGAC	960
QY	961	CTGTGTGGACCACACACCTGTGTACACCCCTGGGCATGTGCTGATGTGGTATCATGTGTGAC	1020
Db	961	CTGTGTGGACCACACACCTGTGTACACCCCTGGGCATGTGCTGATGTGGTATCATGTGTGAC	1020
QY	1021	CCCAAGAGAGCTGCTCTGTCAATGTGAGAGAGATGGGCTTCCATCAGCCCTTACCACTGGC	1080
Db	1021	CCCAAGAGAGCTGCTCTGTCAATGTGAGAGAGATGGGCTTCCATCAGCCCTTACCACTGGC	1080
QY	1081	CACGAGCTGGGGCCACGTTTCAACATGTACCCCAATGTACAAATGTAAAGTCTGTGACGAGGTG	1140
Db	1081	CACGAGCTGGGGCCACGTTTCAACATGTACCCCAATGTACAAATGTAAAGTCTGTGACGAGGTG	1140
QY	1141	TTTGGGAGAGTCCGAGGCAACCAACATGATGTCCCGGACCCCTATCCAGTGTGACGCTGGC	1200
Db	1141	TTTGGGAGAGTCCGAGGCAACCAACATGATGTCCCGGACCCCTATCCAGTGTGACGCTGGC	1200
QY	1201	AAACCCCTGTGCTAGCCTGTGCACTGTGCTGCATCATCACGACTTCTCTGTGACACGCGGACG	1260
Db	1201	AAACCCCTGTGCTAGCCTGTGCACTGTGCTGCATCATCACGACTTCTCTGTGACACGCGGACG	1260
QY	1261	GACTGCTCCTGTGGACCAACCCACAGACCCCATCTCCCTCCCGGAGATCTGCCGGGGCC	1320
Db	1261	GACTGCTCCTGTGGACCAACCCACAGACCCCATCTCCCTCCCGGAGATCTGCCGGGGCC	1320
QY	1321	AGCTACACCCCTGAGCAGCAGTGTGCACTGTGGCTTTTGTGGGCTGTGTCCAAAGCCCTGTCT	1380
Db	1321	AGCTACACCCCTGAGCAGCAGTGTGCACTGTGGCTTTTGTGGGCTGTGTCCAAAGCCCTGTCT	1380
QY	1381	TACATGCACTGTGTGCACCAACGCTGTGTGTGACACCGGAGAGGCCAAGGAGACAGTGTGTGC	1440
Db	1381	TACATGCACTGTGTGCACCAACGCTGTGTGTGACACCGGAGAGGCCAAGGAGACAGTGTGTGC	1440
QY	1441	CAGACCCGCACTTCCCTGTGGGCGGATGTGACACAGCTGTGGGAGAGGCAAGCTGTGCTGC	1500
Db	1441	CAGACCCGCACTTCCCTGTGGGCGGATGTGACACAGCTGTGGGAGAGGCAAGCTGTGCTGC	1500
QY	1501	AAAGGGGCTGTGCTGTGGAGAGACCAACCTTCAACAGACACAGGCTGTGCTGTGCGGC	1560
Db	1501	AAAGGGGCTGTGCTGTGGAGAGACCAACCTTCAACAGACACAGGCTGTGCTGTGCGGC	1560
QY	1561	AAATGGGATCCCTATGGGCCCTCTCTGCGCACATGTGTGTGGGGGCTGTGACGTGTGCAGG	1620
Db	1561	AAATGGGATCCCTATGGGCCCTCTCTGCGCACATGTGTGTGGGGGCTGTGACGTGTGCAGG	1620
QY	1621	AGGCACTGCACCAACCCACCCCTGTGCACACGGGGCAATCTGTGGAGAGGATGTAGGGGTG	1680
Db	1621	AGGCACTGCACCAACCCACCCCTGTGCACACGGGGCAATCTGTGGAGAGGATGTAGGGGTG	1680
QY	1681	AAATACCGATCTGTGCATCTGTGAGACCTGTGCCAGCTCAAGCTTCGGGAAGGGCTTCGGG	1740
Db	1681	AAATACCGATCTGTGCATCTGTGAGACCTGTGCCAGCTCAAGCTTCGGGAAGGGCTTCGGG	1740
QY	1741	GAGGAGCAGTGTAGGCTTTTCAACGGGCTCAACACAGCAGCACACCGGCTCACTTCCGC	1800
Db	1741	GAGGAGCAGTGTAGGCTTTTCAACGGGCTCAACACAGCAGCACACCGGCTCACTTCCGC	1800
QY	1801	GTGGCATGGGTGCCCAAGTACTCTCGGGCTGTCTCCCGGGACAACTGCAAGCTATCTTGC	1860
Db	1801	GTGGCATGGGTGCCCAAGTACTCTCGGGCTGTCTCCCGGGACAAAGTGTCAATCTTGC	1860
QY	1861	CGAGCCATGGCACTGGCTACTCTCTATGTGTGTGTGACACCCAAAGTGTGTGACCGGACGCTG	1920
Db	1861	CGAGCCATGGCACTGGCTACTCTCTATGTGTGTGTGACACCCAAAGTGTGTGACCGGACGCTG	1920
QY	1921	TGCTCTCTGACTGCACCTCCGCTGTGTGTGCACAGGCAAGTGCATCAAGGCTGGCTGTAT	1980

Db	1921	TGCTCTCTACTCTCCACTCCGCTGTGTGTGTCCAAAGGCAAGTCCATCAAGCCTGGCTGTGAT	1980
QY	1981	GGGAACTGGGCTCTCCAAAGAAAGATTCGACAAATGTGTGGGTGTGTGGGGGAGACAAATPAG	2040
Db	1981	GGCAACTCGGGCTCTCAAGAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGGAGACAAATPAG	2040
QY	2041	AGCTGCAAGAAAGGTGACTGGACTCTTCACCAAGGCCAATGATGGCTPACAAATTTTCGTGGTGTG	2100
Db	2041	AGCTGCAAGAAAGGTGACTGGACTCTTCACCAAGGCCAATGATGGCTPACAAATTTTCGTGGTGTG	2100
QY	2101	GCCATCCCGCAGAGCGGCTCTCAAGCANTCGACATCCGCCAGCGCGGTTACAAAGGGCTGATC	2160
Db	2101	GCCATCCCGCAGAGCGGCTCTCAAGCANTCGACATCCGCCAGCGCGGTTACAAAGGGCTGATC	2160
QY	2161	GGGGATGCAAACTACCTGGCTCTGAAGAACAGCCAAAGGCAAGTACCTGTCTCAACGGGCAT	2220
Db	2161	GGGGATGCAAACTACCTGGCTCTGAAGAACAGCCAAAGGCAAGTACCTGTCTCAACGGGCAT	2220
QY	2221	TTCTGTGGTGGCGGGGTGGAGACGGGACCTGGTGGTGAAGGAGCACTGCTGCGCTAACAGC	2280
Db	2221	TTCTGTGGTGGCGGGGTGGAGACGGGACCTGGTGGTGAAGGAGCACTGCTGCGGCTAACAGC	2280
QY	2281	GGCAGCGGGCACACGCGTGGAGAGCCTTGAGAGGCTTCCCGGGCCATCTCTGGAGCGGCTAAC	2340
Db	2281	GGCAGCGGGCACACGCGTGGAGAGCCTTGAGAGGCTTCCCGGGCCATCTCTGGAGCGGCTAAC	2340
QY	2341	GTGAGAGGCTCTCTCGTGGGGGAAATGACACCGCCCGGGGGCGCTACTCTCTTATCTG	2400
Db	2341	GTGAGAGGCTCTCTCGTGGGGGAAATGACACCGCCCGGGGGCGCTACTCTCTTATCTG	2400
QY	2401	CCCAAGAGCCTCGGAGAGAGACAAATCTCTCATCCCAAGAGACCCCGGGGACCCCTCTGTG	2460
Db	2401	CCCAAGAGCCTCGGAGAGAGACAAATCTCTCATCCCAAGAGACCCCGGGGACCCCTCTGTG	2460
QY	2461	TTTGCAACAACAGCTGCTCAAGCCTCTCCAAACAGGTGAGAGACCGCGGAGCGACAGCGCCCT	2520
Db	2461	TTTGCAACAACAGCTGCTCAAGCCTCTCCAAACAGGTGAGAGACCGCGGAGCGACAGCGCCCT	2520
QY	2521	GCACGCTGGGTGGCTGGCAGCTGGGGGCCGTGCTCCGCGAGCCTGAGCAGTGGCCTGAC	2580
Db	2521	GCACGCTGGGTGGCTGGCAGCTGGGGGCCGTGCTCCGCGAGCCTGAGCAGTGGCCTGAC	2580
QY	2581	AAGGGGCGGTGGACTGTGCGGGGCTCCGCGGGGAGAGGAGCAGGCTGCTGTGATGCA	2640
Db	2581	AAGGGGCGGTGGACTGTGCGGGGCTCCGCGGGGAGAGGAGCAGGCTGCTGTGATGCA	2640
QY	2641	GGCCATGGGCGCGTGGAGACAAAGCCTGCGGGGAGAGCCCTGACCTGGAGACTCAGC	2700
Db	2641	GGCCATGGGCGCGTGGAGACAAAGCCTGCGGGGAGAGCCCTGACCTGGAGACTCAGC	2700
QY	2701	GGCTGTGTCACCTCTGCTCCAAAGCTGTGGGCGCGGGGATTTTCAGAGAGGCGCTCACTCAATGT	2760
Db	2701	GGCTGTGTCACCTCTGCTCCAAAGCTGTGGGCGCGGGGATTTTCAGAGAGGCGCTCACTCAATGT	2760
QY	2761	GTGGGCAACGAGAGCGCGGCTGTGGCCGAGGACAGTGTGCAACTTGCACCGCAACGCCAG	2820
Db	2761	GTGGGCAACGAGAGCGCGGCTGTGGCCGAGGACAGTGTGCAACTTGCACCGCAACGCCAG	2820
QY	2821	GAGCTGGACTTCTGCGTCTTGAGGCGCGCTGCA	2883
Db	2821	GAGCTGGACTTCTGCGTCTTGAGGCGCGCTGCA	2883

RESULT 2
 US-09-965-631-7
 ; Sequence 7, Application US/09965631
 ; Patent No. US20020115842A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fridde, Carl Johan
 ; APPLICANT: Hilbun, Erin
 ; TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Enco
 ; FILE REFERENCE: EX-0241-USA

```

? CURRENT APPLICATION NUMBER: US/09/965,631
? CURRENT FILING DATE: 2001-09-27
? PRIOR APPLICATION NUMBER: US 60/236,689
? PRIOR FILING DATE: 2000-09-29
? NUMBER OF SEQ ID NOS: 7
? SOFTWARE: FastSeq for Windows Version 4.0.
? SEQ ID NO 7
? LENGTH: 3446
? TYPE: DNA
? ORGANISM: homo sapiens
? US-09-965-631-7

```

Query match	100.0%;	Score 2853;	DB 10;	Length 3446;
-------------	---------	-------------	--------	--------------

QY	1	ATGTTCTGCGGGGATCTCTAAACCTGGCTTTCCGCGGGGCGAACCGGCGAGGGCTGTAG	60
Db	397	ATGGCTTGTGTGGGATCTTAACCTGGCTTTCCGCGGGGCGAACCGGCTGTGTAG	456
QY	61	CCAGAGCGGAGAGTATGTGTTTCCATCCGACTGACCCGGACATTAAAGCGCGCGCTAC	120
Db	457	CCAGAGCGGAGAGTATGTGTTTCCATCCGACTGACCCGGACATTAAAGCGCGCGCTAC	516
QY	121	TACTTGGCGGGTCCCGAGAGACTCCGGGGATCAGGGACTATTTTCAGATCAGACATT	180
Db	517	TACTTGGCGGGTCCCGAGAGACTCCGGGGATCAGGGACTATTTTCAGATCAGACATT	576
QY	181	CAGGAGGACTTTTACCTACACTGACGCGGATGTCAGTTTGTAGTCCCGCTTCTCC	240
Db	577	CAGGAGGACTTTTACCTACACTGACGCGGATGTCAGTTTGTAGTCCCGCTTCTCC	636
QY	241	ACTGAGCATCTGGGGGTCCCTCCAGGGGGTCCACCGGGGGGTCTTCAGACTGCGAGC	300
Db	637	ACTGAGCATCTGGGGGTCCCTCCAGGGGGTCCACCGGGGGGTCTTCAGACTGCGAGC	696
QY	301	TGCTCTATCTGGGGAGCTGAACCGCGAGCGGACTGTTCGGTCTGTAGACCTGTGC	360
Db	697	TGCTCTATCTGGGGAGCTGAACCGCGAGCGGACTGTTCGGTCTGTAGACCTGTGC	756
QY	361	GGGGGGCTCCGCGGACCTTTGGCTTACCGAGCGCGCGAGTATGTATTAAGCCCGTGC	420
Db	757	GGGGGGCTCCGCGGACCTTTGGCTTACCGAGCGCGCGAGTATGTATTAAGCCCGTGC	816
QY	421	AATGCTAGCGCGCCGGCGCGCGCAAGCGCAACAGCGAGGGCGACACTTCTCAGCGCGG	480
Db	817	AATGCTAGCGCGCCGGCGCGCGCAAGCGCAACAGCGAGGGCGACACTTCTCAGCGCGG	876
QY	481	GGGTGTCCGGGGGGGCTTCGCGAGACCCCACTCTCGCGGGGGTGGCTCGGGCTGG	540
Db	877	GGGTGTCCGGGGGGGCTTCGCGAGACCCCACTCTCGCGGGGGTGGCTCGGGCTGG	936
QY	541	AAACCCGCCATCTTACGGGGCCCTGAGACCTTACAAAGCCGCGGGGGGGCTTCCGGGGAG	600
Db	937	AAACCCGCCATCTTACGGGGCCCTGAGACCTTACAAAGCCGCGGGGGGGCTTCCGGGGAG	996
QY	601	AGTCGTAGCCGGCGAGAGTCTGGGGCGCGCAACGCTTCTGTCTATCCCGGGTACGTG	660
Db	997	AGTCGTAGCCGGCGAGAGTCTGGGGCGCGCAACGCTTCTGTCTATCCCGGGTACGTG	1056
QY	661	GAGACCTGTGTGTCGCGACAGAGTAAATGTGAATTTCCAGCGCGCGGACTGGAACAT	720
Db	1057	GAGACCTGTGTGTCGCGACAGAGTAAATGTGAATTTCCAGCGCGCGGACTGGAACAT	1116
QY	721	TATCTGCTAGCGTGTGTGCAACGCGCGCGCGACTTACCGCATCCAGCATCTCTAAC	780
Db	1117	TATCTGCTAGCGTGTGTGCAACGCGCGCGCGACTTACCGCATCCAGCATCTCTAAC	1176
QY	781	CCCATACATCGTTGTGTGCAAGGTGTGCTCTTAAGATGTAGTACTCCGGGGCAAG	840
Db	1177	CCCATACATCGTTGTGTGCAAGGTGTGCTCTTAAGATGTAGTACTCCGGGGCAAG	1236
QY	841	GTCACCGGGAATCGCGCCCTGACGTGCGCAACTTCTGTGCTCGGAGAAAGAGCTGAAC	900

Db	1237	GTCAACGGCAATGCGGCCCTGAGCGTGGCCAACTTCTGGCTGGCAGAAGAGCTGAAC	1296
Qy	901	AAAGTGAATGACAGCAACCCCCAGTACTGGGAACCTGCCATCTCTTCCACAGGCAGAC	960
Db	1297	AAAGTGAATGACAGCAACCCCCAGTACTGGGAACCTGCCATCTCTTCCACAGGCAGAC	1356
Qy	961	CTGTGTGGAGCCACACACTGTGACCCCTGGACCCCTGGATGGGCTGATGGATCTGTGAC	1020
Db	1357	CTGTGTGGAGCCACACACTGTGACCCCTGGACCCCTGGATGGGCTGATGGATCTGTGAC	1416
Qy	1021	CCCAAGAGAAAGCTGCTCTGTCTTATGAGGAGATGGGGCTTTCATCAGCTTCCACCACTGGC	1080
Db	1417	CCCAAGAGAAAGCTGCTCTGTCTTATGAGGAGATGGGGCTTTCATCAGCTTCCACCACTGGC	1476
Qy	1081	CACGAGCTGGGGCCACGGTTCACATGCCCCCAATGMAATGMAAGTGTGACAGAGCTG	1140
Db	1477	CACGAGCTGGGGCCACGGTTCACATGCCCCCAATGMAATGMAAGTGTGACAGAGCTG	1536
Qy	1141	TTTTGGAGAGCTCCAGGCAACCAACATGATGTCCCGACCCCTATCCAGATGACCGTGGC	1200
Db	1537	TTTTGGAGAGCTCCAGGCAACCAACATGATGTCCCGACCCCTATCCAGATGACCGTGGC	1596
Qy	1201	AAACCCCTGGTACGCTCAGCTCAGTGTGCCATATCACCGAATTCCTGTTGGAAGGGGACAGST	1260
Db	1597	AAACCCCTGGTACGCTCAGTGTGCCATATCACCGAATTCCTGTTGGAAGGGGACAGST	1656
Qy	1261	GACCTGCTCTGTGGACCAACCCAGAACCCATCTCCCTGCCGAGATCTCCCGGGGCC	1320
Db	1657	GACCTGCTCTGTGGACCAACCCAGAACCCATCTCCCTGCCGAGATCTCCCGGGGCC	1716
Qy	1321	AGCTACACCCCTGAGCCACAGATGCGAGCTGGCTTTTGGCGTGGCTCCAAAGCCTGTGCT	1380
Db	1717	AGCTACACCCCTGAGCCACAGATGCGAGCTGGCTTTTGGCGTGGCTCCAAAGCCTGTGCT	1776
Qy	1381	TGCATGACGATCTGACACCAAGCTGTGGTGACCGGGAAAGCCAAAGGACAGATGATGTGC	1440
Db	1777	TGCATGACGATCTGACACCAAGCTGTGGTGACCGGGAAAGCCAAAGGACAGATGATGTGC	1836
Qy	1441	CAGACCCGCCCACTTCCCTGGGGCCGATGAGCAGCAGCTGTGGCGAGGGCAAGCTGTGCC	1500
Db	1837	CAGACCCGCCCACTTCCCTGGGGCCGATGAGCAGCAGCTGTGGCGAGGGCAAGCTGTGCC	1896
Qy	1501	AAAGGGGCTCGGTGGAGAGACAAACCTCAAAACCAAGGGGTGGATGGTCTCGGGCC	1560
Db	1897	AAAGGGGCTCGGTGGAGAGACAAACCTCAAAACCAAGGGGTGGATGGTCTCGGGCC	1956
Qy	1561	AAATGGGATTCCTTATGAGCCCTGTCTCCGACATGTGTGGGGGGGTGTGAGCTGTGGCAGG	1620
Db	1957	AAATGGGATTCCTTATGAGCCCTGTCTCCGACATGTGTGGGGGGGTGTGAGCTGTGGCAGG	2016
Qy	1621	AGGCAATGACCAACCCACCCCTGTCCCAAGGGGGGCAAGTACTGCGAAGGAGATGAGGTG	1680
Db	2017	AGGCAATGACCAACCCACCCCTGTCCCAAGGGGGGCAAGTACTGCGAAGGAGATGAGGTG	2076
Qy	1681	AAATACCGATCTCGAATCTGGACCCCTGTGCCAGCTCAAGCTCCGGAAGAAGCTTCCGG	1740
Db	2077	AAATACCGATCTCGAATCTGGACCCCTGTGCCAGCTCAAGCTCCGGAAGAAGCTTCCGG	2136
Qy	1741	GAGGAGCACTGTGAGGCTTTTCAAGGCTACACCAACCAAGCCACCAACCCGCTCACTTCCCC	1800
Db	2137	GAGGAGCACTGTGAGGCTTTTCAAGGCTACACCAACCAAGCCACCAACCCGCTCACTTCCCC	2196
Qy	1801	GTTCGATGGGTGCCCAATACTCCGGGGGTGTCTCCCGGGACCAAGTCAAGTCAATCTGC	1860
Db	2197	GTTCGATGGGTGCCCAATACTCCGGGGGTGTCTCCCGGGACCAAGTCAAGTCAATCTGC	2256
Qy	1861	CGAGCCAAATGAGCACTGGCTACTTCTATGTGCTGGCAACCAAGTGTGAGAGGCAAGCTG	1920
Db	2257	CGAGCCAAATGAGCACTGGCTACTTCTATGTGCTGGCAACCAAGTGTGAGAGGCAAGCTG	2316
Qy	1921	TGCTCTCTGACTCACTCCGTCTGTGTCCAAAGCAAGTCAAGCTGGCTGTAT	1980

Db 2317 TGTCTCTGATCCACCTCCGCTGTGTCTCCAGGCAAGTGCATCAAGCTGGCTGTAT 2376
QY 1981 GGGAACTGGGCTCCAGAAAGAGATTGCAGAGTGTGGGGTGTGGGGGAGNAAATAG 2040
Db 2377 GGGAACTGGGCTCCAGAAAGATTCAGAAAGTGTGGGGTGTGGGGGAGNAAATAG 2436
QY 2041 AGCTGCAAGAAAGTACTGACTGTTCACCAAGCCATGCTACAAATTTCTGTGTG 2100
Db 2437 AGCTGCAAGAAAGTACTGACTGTTCACCAAGCCATGCTACAAATTTCTGTGTG 2496
QY 2101 GGCATCCCCGAGGCGCTCAAGCATGACATCCGCCAGGCGGTTCAGAAAGGCTGATC 2160
Db 2497 GGCATCCCCGAGGCGCTCAAGCATGACATCCGCCAGGCGGTTCAGAAAGGCTGATC 2556
QY 2161 GGGGATGCAACTACTGTGCTCTGAAGACAGCCAGCAAGTACTGTCTCAAGGCGCAT 2220
Db 2557 GGGGATGCAACTACTGTGCTCTGAAGACAGCCAGCAAGTACTGTCTCAAGGCGCAT 2516
QY 2221 TTGCTGTGTGGGGGTGGAGCGGGACCTGTGTGTAAGGCGCATGCTGTGGGTACAGC 2280
Db 2617 TTGCTGTGTGGGGGTGGAGCGGGACCTGTGTGTAAGGCGCATGCTGTGGGTACAGC 2676
QY 2281 GGCACGGGACAGCGGTGGAGAGCGCTGACAGCTCCCGGCCCATCTGGAGCGGCTGACC 2340
Db 2677 GGCACGGGACAGCGGTGGAGAGCGCTGACAGCTCCCGGCCCATCTGGAGCGGCTGACC 2736
QY 2341 GTGAGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2737 GTGAGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2796
QY 2401 CCCAAAGGCTCTGGGGAAGCAAGTCTCTCATTCCTCAAGAGACCCCGGGGAGCCCTCTGTG 2460
Db 2797 CCCAAAGGCTCTGGGGAAGCAAGTCTCTCATTCCTCAAGAGACCCCGGGGAGCCCTCTGTG 2856
QY 2461 TTGCACAACAGCGTCTCTCAGCTCTCCAAACAGGTGGAGAGCGGACAGAGCCCCCT 2520
Db 2857 TTGCACAACAGCGTCTCTCAGCTCTCCAAACAGGTGGAGAGCGGACAGAGCCCCCT 2916
QY 2521 GCACGCTGGGTGGCTGGCACTGGGGGCGCTGCTCCGCGAGCTGGGCGAGTGGCTTCAG 2580
Db 2917 GCACGCTGGGTGGCTGGCACTGGGGGCGCTGCTCCGCGAGCTGGGCGAGTGGCTTCAG 2976
QY 2581 AAGCGGCGGTGGAGTCTCGGGGCTCGCGGCGAGCGGAGCGGCTCTCTGTGTATGTA 2640
Db 2977 AAGCGGCGGTGGAGTCTCGGGGCTCGCGGCGAGCGGAGCGGCTCTCTGTGTATGTA 3036
QY 2641 GCCCATCGGCGGTGGAGACAAAGCTTGGGGGAGCCCTGCCCCACCTTGGAGCTCAGC 2700
Db 3037 GCCCATCGGCGGTGGAGACAAAGCTTGGGGGAGCCCTGCCCCACCTTGGAGCTCAGC 3096
QY 2701 GCGTGTGCAACCTGTCTCCAGAGCTGGGCGCGGGATTTCAGAGCGCTCAAGTGT 2760
Db 3097 GCGTGTGCAACCTGTCTCCAGAGCTGGGCGCGGGATTTCAGAGCGCTCAAGTGT 3156
QY 2761 GTGGGCGACAGAGCGGCTGTGGGCGCGGAGCAAGTGCAGCTTGCACGCAAGGCCGAG 2820
Db 3157 GTGGGCGACAGAGCGGCTGTGGGCGCGGAGCAAGTGCAGCTTGCACGCAAGGCCGAG 3216
QY 2821 GAGCTGACTTCTGCTCTGAGGCGGCTGTA 2853
Db 3217 GAGCTGACTTCTGCTCTGAGGCGGCTGTA 3249

RESULT 3

US-10-163-316-3
; Sequence 3, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MP101-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316

;; CURRENT FILING DATE: 2002-06-05
;; PRIOR APPLICATION NUMBER: 60/297,863
;; PRIOR FILING DATE: 2001-06-13
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 2469
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(2469)
US-10-163-316-3
Query Match 52.3%; Score 1491; DB 9; Length 2469;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTTCTGCTGGGATCCTTAACCTGGCTTTCGCGGGGCAACCGCTGGAGGCTGTAG 60
Db 1 ATGCTTCTGCTGGGATCCTTAACCTGGCTTTCGCGGGGCAACCGCTGGAGGCTGTAG 60
QY 61 CCAGAGCGGAGGTAGTGTCCATCCGACTGACCCGGAATTACGGCGCGCTAC 120
Db 61 CCAGAGCGGAGGTAGTGTCCATCCGACTGACCCGGAATTACGGCGCGCTAC 120
QY 121 TACTGGCGGGGTCGCGAGAGTCCGGGATCAGAGCACTATTTTCAATACAGCATTT 180
Db 121 TACTGGCGGGGTCGCGAGAGTCCGGGATCAGAGCACTATTTTCAATACAGCATTT 180
QY 181 CAGAGAGACTTTTACCTACCTGACCGCGGATGCTAGTTCCTTGGCTCCGCTTCC 240
Db 181 CAGAGAGACTTTTACCTACCTGACCGCGGATGCTAGTTCCTTGGCTCCGCTTCC 240
QY 241 ACTAGCATCTGGGGTCCCTCCAGAGGCTCACCGGGGCTTTACAGCTTGCAGC 300
Db 241 ACTAGCATCTGGGGTCCCTCCAGAGGCTCACCGGGGCTTTACAGCTTGCAGC 300
QY 301 TGCTTCTATTTCTGGGAGGTGACAGCGGAGCGGACTGTTCGGTGTGAGCTGTGC 360
Db 301 TGCTTCTATTTCTGGGAGGTGACAGCGGAGCGGACTGTTCGGTGTGAGCTGTGC 360
QY 361 GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTATTAAGCCGCTCCC 420
Db 361 GGGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGAGTATGTATTAAGCCGCTCCC 420
QY 421 AATGTACGCGGCGGCGGCGGAGCAAGAGCGGAGCGGACCTTTCACAGCGCGG 480
Db 421 AATGTACGCGGCGGCGGCGGAGCAAGAGCGGAGCGGACCTTTCACAGCGCGG 480
QY 481 GGTGTTCGCGGCGGCTTCCGAGACCCACCTCTGTGCGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCGCGGCGGCTTCCGAGACCCACCTCTGTGCGGGGTGGCTCGGGCTGG 540
QY 541 AACCCTGCTATCTACAGGCGCTTGAGCCCTTACAGCCGCGGGGCTTGGGGAG 600
Db 541 AACCCTGCTATCTACAGGCGCTTGAGCCCTTACAGCCGCGGGGCTTGGGGAG 600
QY 601 AGTGTACCGCGGCGGAGTGTGGGCGGCAAGGCTTGTGCTATCCGCGGTTACG 660
Db 601 AGTGTACCGCGGCGGAGTGTGGGCGGCAAGGCTTGTGCTATCCGCGGTTACG 660
QY 661 GAGAGCGTGTGTGCGGAGAGTCAATGATTCACAGGCGGAGCTGTGAACAT 720
Db 661 GAGAGCGTGTGTGCGGAGAGTCAATGATTCACAGGCGGAGCTGTGAACAT 720
QY 721 TATCTGTGAGGCTGTGCGCAAGCGGCGGAGCTTACCGCATCCAGACTCTCAAC 780
Db 721 TATCTGTGAGGCTGTGCGCAAGCGGCGGAGCTTACCGCATCCAGACTCTCAAC 780
QY 781 CCCATCAACATCGTTGTGCAAGGTGCTGCTTGTGAGATGTGATCCGGGGCCCAAG 840
Db 781 CCCATCAACATCGTTGTGCAAGGTGCTGCTTGTGAGATGTGATCCGGGGCCCAAG 840

Qy	841	GTCCCGGGCAATGCGGGCCCTGACGCGTGGGCACTCTGCGCTGGGCAAGAAACACTTAAC	900
Qy	841	GTCCCGGGCAATGCGGGCCCTGACGCGTGGGCACTCTGCGCTGGGCAAGAAACACTTAAC	900
Qy	841	GTCCCGGGCAATGCGGGCCCTGACGCGTGGGCACTCTGCGCTGGGCAAGAAACACTTAAC	900
Qy	901	AAAGTAGTACAAAGCAACCCGAGTACTGGGCACTGGCAATCTCTTCACCAAGGCAAGAC	960
Db	901	AAAGTAGTACAAAGCAACCCGAGTACTGGGCACTGGCAATCTCTTCACCAAGGCAAGAC	960
Qy	961	CTGTGTGGAGACCACCACTGTGTACACCCCTGGGCAATGGCTGATGTGGGTACCATGTGTGAC	1020
Db	961	CTGTGTGGAGCCCAACACTGTGTGACCCCTGGGCAATGGCTGATGTGGGTACCATGTGTGAC	1020
Qy	1021	CCCAAGAGAAAGCTGCTCTGTCAATTAGAGGACGATGGGGCTTCCATCAGCCCTTCACCACTGCC	1080
Db	1021	CCCAAGAGAAAGCTGCTCTGTCAATTAGAGGACGATGGGGCTTCCATCAGCCCTTCACCACTGCC	1080
Qy	1081	CACGAGCTGGGGCCACAGTGTTCACATGATCCCCCATGTACAAATGTAAAGTGTGTAGAGAGGTG	1140
Db	1081	CACGAGCTGGGGCCACAGTGTTCACATGATCCCCCATGTACAAATGTAAAGTGTGTAGAGAGGTG	1140
Qy	1141	TTTGGGAAGCTCCGAGCAACCAATGATGTCCCGACCCCTATCCAGATGACAGCGTGTGC	1200
Db	1141	TTTGGGAAGCTCCGAGCAACCAATGATGTCCCGACCCCTATCCAGATGACAGCGTGTGC	1200
Qy	1201	AACCCCTGTGACGCTGTGACAGTGTGCAATACCGACTTCTTSGACAGCGGGCAAGGT	1260
Db	1201	AACCCCTGTGACAGCTGTGACAGTGTGCAATACCGACTTCTTSGACAGCGGGCAAGGT	1260
Qy	1261	GACTGCTCTCCCTGGGCAACCCACACAGGCCATCTCCCTGCGGAGSAGATCTGGCGGGGCGCC	1320
Db	1261	GACTGCTCTCCCTGGGCAACCCACACAGGCCATCTCCCTGCGGAGSAGATCTGGCGGGGCGCC	1320
Qy	1321	AGCTACACCCCTGAGCCAGCAGTGGAGTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCT	1380
Db	1321	AGCTACACCCCTGAGCCAGCAGTGGAGTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCT	1380
Qy	1381	TACATGCAATACACCAAGAGTGTGTGACACCGGGAAGGCCAAGAGGACATATGTTGTGC	1440
Db	1381	TACATGCAATACACCAAGAGTGTGTGACACCGGGAAGGCCAAGAGGACATATGTTGTGC	1440
Qy	1441	CAGACCCGCGCACTTCCCTGTGGGCGGATGTGGACACAGCTGTGGGAGAGGCAAGCTGTGCCCT	1500
Db	1441	CAGACCCGCGCACTTCCCTGTGGGCGGATGTGGACACAGCTGTGGGAGAGGCAAGCTGTGCCCT	1500
Qy	1501	AAAGGGGCTGTGCTGGAGAGAACACAACTTCACACAGCAAG	1542
Db	1501	AAAGGGGCTGTGCTGGAGAGAACACAACTTCACAAACACAGG	1542

```

RESULT 4
US-10-163-316-1
; Sequence 1, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP101-025P11NM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:
;

```

NAME/KEY	misc. feature	LOCATION	(1)...	(2940)
OTHER INFORMATION:	n = A,T,C or G			
US-10-163-316-1				
Query Match	52.3%;	Score 1491;	DB 9;	Length 2940;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1541;	Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;
QY	1	ATGCTTTCGTCTGGCATCTTAACCTTGCTTTCGCCGGCGGAACCGCTGAGAGCTCTGAG	60	
DB	472	ATGCTTTCGTCTGGCATCTTAACCTTGCTTTCGCCGGCGGAACCGCTGAGAGCTCTGAG	531	
QY	61	CCAGAGCGGGAGGTAGTGTCTTCCATCCGACATGAGACCCGGACATTTAAAGCGCGCGCTAC	120	
DB	532	CCAGAGCGGGAGGTAGTGTCTTCCATCCGACATGAGACCCGGACATTTAAAGCGCGCGCTAC	591	
QY	121	TACTGGCGGGGTCCCGAGAGCTCCGGGGATCAGGGACTATTTTTCAGATCAGACGAT	180	
DB	592	TACTGGCGGGGTCCCGAGAGCTCCGGGGATCAGGGACTATTTTTCAGATCAGACGAT	651	
QY	181	CAGGAGAGATTTTACTTAACCTGAGACCGCGGATGCTCATTTCTGGCTCCCGCTCTCTC	240	
DB	652	CAGGAGAGATTTTACTTAACCTGAGACCGCGGATGCTCATTTCTGGCTCCCGCTCTCTC	711	
QY	241	ACTGACATCTGAGGCGTCCCGCTCCAGGGGGTACACGGGGGCTCTTCAACACGCGACAC	300	
DB	712	ACTGACATCTGAGGCGTCCCGCTCCAGGGGGTACACGGGGGCTCTTCAACACGCGACAC	771	
QY	301	TGCTTATTTCTGGGGAGCTGAACCGCGAGCGGACTCTTGCTGTGAGCTGTGCTG	360	
DB	772	TGCTTATTTCTGGGGAGCTGAACCGCGAGCGGACTCTTGCTGTGAGCTGTGCTG	831	
QY	361	GGGGGGCTCCGGGAGACCTTTGGCTACCGAGGGCGCGAGATGATCATTTAGCCCGCGCCC	420	
DB	832	GGGGGGCTCCGGGAGACCTTTGGCTACCGAGGGCGCGAGATGATCATTTAGCCCGCGCCC	891	
QY	421	AATGCTAGCGCGCGCGCGCGCGACGACGACGACGACGACGACGACGACGACGACGACG	480	
DB	892	AATGCTAGCGCGCGCGCGCGCGACGACGACGACGACGACGACGACGACGACGACGACG	951	
QY	481	GGTGTTCGGGGCGGGGCTTCCGGAGACCCCACTCTGCTGCGGGGTGGCTCTGGGCTG	540	
DB	952	GGTGTTCGGGGCGGGGCTTCCGGAGACCCCACTCTGCTGCTGCGGGGTGGCTCTGGGCTG	1011	
QY	541	AACCCCGCATCTTAAGGGGCGCTTGAGACCTTAACCGCGGGGGCGGGGCTTCCGGGAG	600	
DB	1012	AACCCCGCATCTTAAGGGGCGCTTGAGACCTTAACCGCGGGGGCGGGGCTTCCGGGAG	1071	
QY	601	AGTGTAGCGCGCGCGAGGCTTGGGGCGCGCACGCTTGTCTGTATCCCGCGGTACGT	660	
DB	1072	AGTGTAGCGCGCGCGAGGCTTGGGGCGCGCACGCTTGTCTGTATCCCGCGGTACGT	1131	
QY	661	GAGAGCTGTGTGTCGCGAGAGTCAATGTCAAGTTTCCAGCGCGCGGACCTTGAGACAT	720	
DB	1132	GAGAGCTGTGTGTCGCGAGAGTCAATGTCAAGTTTCCAGCGCGCGGACCTTGAGACAT	1191	
QY	721	TATCTGCTAGCGTCTGTGCAAGGGGGCGCGACTTAACCGCATCCCAACATCTCTAAC	780	
DB	1192	TATCTGCTAGCGTCTGTGCAAGGGGGCGCGACTTAACCGCATCCCAACATCTCTAAC	1251	
QY	781	CCCATCAACATCTTGTGTGCAAGGTGCTCTTAAAGATGCTGACTCCGGGCCAAC	840	
DB	1252	CCCATCAACATCTTGTGTGCAAGGTGCTCTTAAAGATGCTGACTCCGGGCCAAC	1311	
QY	841	GTCACCGCGCAATCGGCGCTGACGCTGCGCACTTGTGTGCTGCGAGAAAGCTGAAC	900	
DB	1312	GTCACCGCGCAATCGGCGCTGACGCTGCGCACTTGTGTGCTGCGAGAAAGCTGAAC	1371	
QY	901	AAAGTGTAGTACAAAGACCCCGAGTACTGGGAGACATGCTCTTCAACAGAGAC	960	
DB	1372	AAAGTGTAGTACAAAGACCCCGAGTACTGGGAGACATGCTCTTCAACAGAGAC	1431	

```
QY 961 CTGTGTGAGGACACACCTGTGTGACACCTGGGCGATGGCTGATGTGGGTACCATGTGTGAC 1020
Db 1422 CTGTGTGAGGACACACCTGTGTGACACCTGGGCGATGGCTGATGTGGGTACCATGTGTGAC 1491
QY 1021 CCCAAGAGAGAGCTGTGTGATTTGAGGAGCATGGGCTTCATGAGCCTTCACAGAGGCC 1080
Db 1492 CCCAAGAGAGAGCTGTGTGATTTGAGGAGCATGGGCTTCATGAGCCTTCACAGAGGCC 1551
QY 1081 CACGAGCTGGGCGACAGTGTGTACATGGCCCATGACATGTGAAGTCTGTGAGAGGTG 1140
Db 1552 CACGAGCTGGGCGACAGTGTGTACATGGCCCATGACATGTGAAGTCTGTGAGAGGTG 1611
QY 1141 TTTGGAGAGCTCCGAGCCCAACCAATATGTGTCCCGAGCCTTCATCCAGATGACCGTGGC 1200
Db 1612 TTTGGAGAGCTCCGAGCCCAACCAATATGTGTCCCGAGCCTTCATCCAGATGACCGTGGC 1671
QY 1201 AACCCCTGTGAGCCTGTGAGTGTGATCATCCAGCCTTCCTGGAGAGGGGCGACGGT 1260
Db 1672 AACCCCTGTGAGCCTGTGAGTGTGATCATCCAGCCTTCCTGGAGAGGGGCGACGGT 1731
QY 1261 GACTGCTCTGTGAGCAACCAACGAGAGGCCATCTCCCTGCGGAGATCTGCGGGCGCC 1320
Db 1732 GACTGCTCTGTGAGCAACCAACGAGAGGCCATCTCCCTGCGGAGATCTGCGGGCGCC 1791
QY 1321 AGCTACACCCCTGAGCGAGAGTGGAGAGTGGCTTTGGCGTGGGCTCCAGACCCCTGTCT 1380
Db 1792 AGCTACACCCCTGAGCGAGAGTGGAGAGTGGCTTTGGCGTGGGCTCCAGACCCCTGTCT 1851
QY 1381 TACATGACAGTACTGACCAACAGCTGTGTGACACCGGAGAGGCCAAGSAGACAGATGTGTG 1440
Db 1852 TACATGACAGTACTGACCAACAGCTGTGTGACACCGGAGAGGCCAAGSAGACAGATGTGTG 1911
QY 1441 CAGACCCGCGCATCTCCCTGGGCGGATGTGACACCAAGCTGTGTGCGAGGGCAAGCTGTGCTC 1500
Db 1912 CAGACCCGCGCATCTCCCTGGGCGGATGTGACACCAAGCTGTGTGCGAGGGCAAGCTGTGCTC 1971
QY 1501 AAAGGGGCTGCTGTGAGAGACACACCTCAACAGCAGAG 1542
Db 1972 AAAGGGGCTGCTGTGAGAGACACACCTCAACAGCAGAG 2013

RESULT 5
US-09-965-631-5
; Sequence 5, Application US/09965631
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Fildiddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Encodit
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-5

Query Match 38.2%; Score 1091; DB 10; Length 1104;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 121 TACTGGCGGGGTCGCCGAGAGCTCCGGGGATCAGGGACTCATTTTCAGATCAGACATTT 180
Db 121 TACTGGCGGGGTCGCCGAGAGCTCCGGGGATCAGGGACTCATTTTCAGATCAGACATTT 180
QY 181 CAGGAGAGCTTTTACCTACACCTGAGCAGCGGAGATGTCTAGTTTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGAGCTTTTACCTACACCTGAGCAGCGGAGATGTCTAGTTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGAGGGCTTCACCGGGGGCTTTTCAGACCTGGCAGCC 300
Db 241 ACTGAGCATCTGGGCGTCCCTCCAGAGGGCTTCACCGGGGGCTTTTCAGACCTGGCAGCC 300
QY 301 TGGCTTATTTCTGGGAGACTGAAAGCGCGAGCGGAGCTGTTGGCTGTGTGAGCCTGTGC 360
Db 301 TGGCTTATTTCTGGGAGACTGAAAGCGCGAGCGGAGCTGTTGGCTGTGTGAGCCTGTGC 360
QY 361 GGGGGGCTCCCGGAGACCTTTTGGCTACCGAGGCGCGAGTATGTATTAGCCCGCTGCC 420
Db 361 GGGGGGCTCCCGGAGACCTTTTGGCTACCGAGGCGCGAGTATGTATTAGCCCGCTGCC 420
QY 421 AATGCTAGGCGCGCGGCGGAGGCGAAGGCGAAGGCGGAGGCGGACACTTCTCCAGGCGCG 480
Db 421 AATGCTAGGCGCGCGGCGGAGGCGAAGGCGAAGGCGGAGGCGGACACTTCTCCAGGCGCG 480
QY 481 GGTGTTCGAGGCGGCGCTTCCGAGAGACCCACACTCTCGCTGGGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCGAGGCGGCGCTTCCGAGAGACCCACACTCTCGCTGGGGGGTGGCTCGGGCTGG 540
QY 541 AACCCCGCATCTCTACAGGGGCTTCGACCTTTCAAGCCGCGGCGGGGCTTGGGGAG 600
Db 541 AACCCCGCATCTCTACAGGGGCTTCGACCTTTCAAGCCGCGGCGGGGCTTGGGGAG 600
QY 601 AGTCGTAGCGCGGCGAGGCTGTGGGCGCGCAAGGCTTTGATGTCATCCGCGGAGTG 660
Db 601 AGTCGTAGCGCGGCGAGGCTGTGGGCGCGCAAGGCTTTGATGTCATCCGCGGAGTG 660
QY 661 GAGAGGCTGTGTGTGCGGAGAGTCAATGTCAAGTTTCAAGGCGCGGAGCTGGAAAT 720
Db 661 GAGAGGCTGTGTGTGCGGAGAGTCAATGTCAAGTTTCAAGGCGCGGAGCTGGAAAT 720
QY 721 TATCTGTGAGCGCTGTGGCAAGGCGGCGGAGCTCTACCGCATCCAGATCCTCAAC 780
Db 721 TATCTGTGAGCGCTGTGGCAAGGCGGCGGAGCTCTACCGCATCCAGATCCTCAAC 780
QY 781 CCCATCAACATCTGTGTGTGATCAAGGTGCTGCTTTAGAGATCGTAGCTCGGGCCCAAG 840
Db 781 CCCATCAACATCTGTGTGTGATCAAGGTGCTGCTTTAGAGATCGTAGCTCGGGCCCAAG 840
QY 841 GTACACCGGCAATGGCGCCCTGACGCTGCGCAACTTGTGTGCTGGCAGAGAGAGCTGAAAC 900
Db 841 GTACACCGGCAATGGCGCCCTGACGCTGCGCAACTTGTGTGCTGGCAGAGAGAGCTGAAAC 900
QY 901 AAAGTGTGAGCAAGCAGCCGAGAGTGTGGAGCACTGCAATCCATCCATCCAGGAGAGAG 960
Db 901 AAAGTGTGAGCAAGCAGCCGAGAGTGTGGAGCACTGCAATCCATCCATCCAGGAGAGAG 960
QY 961 CTGTGTGAGGACACCACTGTGTGACACCTGTGGAGATGGCTGTGATGTGGTATCATGTGTGAC 1020
Db 961 CTGTGTGAGGACACCACTGTGTGACACCTGTGGAGATGGCTGTGATGTGGTATCATGTGTGAC 1020
QY 1021 CCCAAGAGAGAGCTGTGTGATTTGAGGAGCATGGGCTTCATGAGCCTTCACAGAGGCC 1080
Db 1021 CCCAAGAGAGAGCTGTGTGATTTGAGGAGCATGGGCTTCATGAGCCTTCACAGAGGCC 1080
QY 1081 CACGAGCTGGG 1091
Db 1081 CACGAGCTGGG 1091

RESULT 6
US-09-965-631-1
; Sequence 1, Application US/09965631
```



```
; Patent No. US20020115842A1
; GENERAL INFORMATION:
; APPLICANT: Fildale, Carl Johan
; TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US-09/965,631
; PRIOR FILING DATE: 2001-09-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 966
; TYPE: DNA
; ORGANISM: homo sapiens
; US-09-965-631-1

Query Match      33.6%; Score 959; DB 10; Length 966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCTTCTGCTGGGACATCTTAACCCCTTGTGGCCGGGGAACCCGTGAGCTCTGAG 60
      1 ATGCTTCTGCTGGGACATCTTAACCCCTTGTGGCCGGGGAACCCGTGAG 60
Db
QY      61 CCAGAGCGGAGGAGTAGTCTTCCCATCCAGCTGAGCCCGGACATTAACGCGCGCCCTGAC 120
      61 CCAGAGCGGAGGAGTAGTCTTCCCATCCAGCTGAGCCCGGACATTAACGCGCGCCCTGAC 120
Db
QY      121 TACTGGCGGGGTCGCCAGAGACTCCGGGAGTCAAGGAGCTATTTTCAGATCACAGCATTT 180
      121 TACTGGCGGGGTCGCCAGAGACTCCGGGAGTCAAGGAGCTATTTTCAGATCACAGCATTT 180
Db
QY      181 CAGGAGGACTTTTACCTACACCTGAGCCCGGAGTCTAGTCTTGGCTCCCGCTCTCC 240
      181 CAGGAGGACTTTTACCTACACCTGAGCCCGGAGTCTAGTCTTGGCTCCCGCTCTCC 240
Db
QY      241 ACTGAGCATCTGGGCTCCCTCCCTCAAGGGGCTCACCGGGGGCTTTTCAGACCTGCGAGCG 300
      241 ACTGAGCATCTGGGCTCCCTCCCTCAAGGGGCTCACCGGGGGCTTTTCAGACCTGCGAGCG 300
Db
QY      301 TGCTTCTATTCTGGGAGCGTAACCGCGGAGCCGAGCTCTGCTGCTGTGAGCTGTGC 360
      301 TGCTTCTATTCTGGGAGCGTAACCGCGGAGCCGAGCTCTGCTGCTGTGAGCTGTGC 360
Db
QY      361 GGGGGGCTCCGGGAGCTTTTGGCTACCGAGGCGCGAGTATGTCAATTAGCCCGCTGCC 420
      361 GGGGGGCTCCGGGAGCTTTTGGCTACCGAGGCGCGAGTATGTCAATTAGCCCGCTGCC 420
Db
QY      421 AATGCTAGCGCGCGCGGCGGCAAGCGCAACAGCGGCGGCAACACCTTCACAGCGCGCG 480
      421 AATGCTAGCGCGCGCGGCGGCAAGCGCAACAGCGGCGGCAACACCTTCACAGCGCGCG 480
Db
QY      481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCTCGGGCTGG 540
      481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCTCGGGCTGG 540
Db
QY      541 AACCCCGCATCTCTAGCGGCGCTTGGACCTTACAAAGCCCGGCGGCGCTTCGCGGAG 600
      541 AACCCCGCATCTCTAGCGGCGCTTGGACCTTACAAAGCCCGGCGGCGCTTCGCGGAG 600
Db
QY      601 AGTCGTAGCGCGCGAGGTCTTGGCGCGCAAGCGCTTCTGTATCCCGGCTACG 660
      601 AGTCGTAGCGCGCGAGGTCTTGGCGCGCAAGCGCTTCTGTATCCCGGCTACG 660
Db
QY      661 GAGACGCTGTGTGCGGAGAGTCAATGTGCAAGTTCCAGGGCGGAGCTTGAACAT 720
      661 GAGACGCTGTGTGCGGAGAGTCAATGTGCAAGTTCCAGGGCGGAGCTTGAACAT 720
Db
QY      721 TATCTGTGACGCTCTGCGCAAGCGGCGGAGCTCTACCGGCATCCAGATCTCTAAC 780
      721 TATCTGTGACGCTCTGCGCAAGCGGCGGAGCTCTACCGGCATCCAGATCTCTAAC 780
Db
```

```
QY      781 CCATCAACATGTTGTGTCGAAGTGTCTTCTTAGAGATCGTGAATCCGGGCCCAAG 840
      781 CCATCAACATGTTGTGTCGAAGTGTCTTCTTAGAGATCGTGAATCCGGGCCCAAG 840
Db
QY      841 GTACCGGCAATGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAAGAGCTGAAC 900
      841 GTACCGGCAATGCGGCCCTGACGCTGCGCACTTCTGTGCTGGCAGAGAAGAGCTGAAC 900
Db
QY      901 AAAGTAGTGACAGACACCCGAGTACTGGGAGCACTGCCATCTCTTACCGAGGACAGA 959
      901 AAAGTAGTGACAGACACCCGAGTACTGGGAGCACTGCCATCTCTTACCGAGGACAGA 959
Db
```

Search completed: May 16, 2003, 10:21:28
Job time : 4157 secs

GenCore version 5.1.4, p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 01:01:57 ; Search time 5576 Seconds
(without alignments)
12864.332 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853
Sequence: 1 atgctctgctggagctcct.....gcgtcttgagcgctgctga 2853

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 24791104 seqs, 12571243825 residues

Word size : 24

Total number of hits satisfying chosen parameters: 123

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

Pending_Patents_NA_Main:*
1: /cgn2_6/prodata/1/pna/US06000_COMB.seq:*
2: /cgn2_6/prodata/1/pna/US06001_COMB.seq:*
3: /cgn2_6/prodata/1/pna/US07000_COMB.seq:*
4: /cgn2_6/prodata/1/pna/US08000_COMB.seq:*
5: /cgn2_6/prodata/1/pna/US08001_COMB.seq:*
6: /cgn2_6/prodata/1/pna/US08002_COMB.seq:*
7: /cgn2_6/prodata/1/pna/US08003_COMB.seq:*
8: /cgn2_6/prodata/1/pna/US08004_COMB.seq:*
9: /cgn2_6/prodata/1/pna/US08005_COMB.seq:*
10: /cgn2_6/prodata/1/pna/US08006_COMB.seq:*
11: /cgn2_6/prodata/1/pna/US08007_COMB.seq:*
12: /cgn2_6/prodata/1/pna/US08008_COMB.seq:*
13: /cgn2_6/prodata/1/pna/US08009_COMB.seq:*
14: /cgn2_6/prodata/1/pna/US09000_COMB.seq:*
15: /cgn2_6/prodata/1/pna/US09001_COMB.seq:*
16: /cgn2_6/prodata/1/pna/US09002_COMB.seq:*
17: /cgn2_6/prodata/1/pna/US09003_COMB.seq:*
18: /cgn2_6/prodata/1/pna/US09004_COMB.seq:*
19: /cgn2_6/prodata/1/pna/US09005_COMB.seq:*
20: /cgn2_6/prodata/1/pna/US09006_COMB.seq:*
21: /cgn2_6/prodata/1/pna/US09007_COMB.seq:*
22: /cgn2_6/prodata/1/pna/US09008_COMB.seq:*
23: /cgn2_6/prodata/1/pna/US09009_COMB.seq:*
24: /cgn2_6/prodata/1/pna/US09010_COMB.seq:*
25: /cgn2_6/prodata/1/pna/US09011_COMB.seq:*
26: /cgn2_6/prodata/1/pna/US09012_COMB.seq:*
27: /cgn2_6/prodata/1/pna/US09013_COMB.seq:*
28: /cgn2_6/prodata/1/pna/US09014_COMB.seq:*
29: /cgn2_6/prodata/1/pna/US09015_COMB.seq:*
30: /cgn2_6/prodata/1/pna/US09016_COMB.seq:*
31: /cgn2_6/prodata/1/pna/US09017_COMB.seq:*
32: /cgn2_6/prodata/1/pna/US09018_COMB.seq:*
33: /cgn2_6/prodata/1/pna/US09019_COMB.seq:*
34: /cgn2_6/prodata/1/pna/US09020_COMB.seq:*
35: /cgn2_6/prodata/1/pna/US09021_COMB.seq:*
36: /cgn2_6/prodata/1/pna/US09022_COMB.seq:*
37: /cgn2_6/prodata/1/pna/US09023_COMB.seq:*
38: /cgn2_6/prodata/1/pna/US09024_COMB.seq:*
39: /cgn2_6/prodata/1/pna/US10000_COMB.seq:*
40: /cgn2_6/prodata/1/pna/US10001_COMB.seq:*
41: /cgn2_6/prodata/1/pna/US10002_COMB.seq:*
42: /cgn2_6/prodata/1/pna/US10003_COMB.seq:*
43: /cgn2_6/prodata/1/pna/US10004_COMB.seq:*

```

```

44: /cgn2_6/prodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/prodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/prodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/prodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/prodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/prodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/prodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/prodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/prodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/prodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/prodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/prodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/prodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/prodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/prodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/prodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/prodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/prodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/prodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/prodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/prodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/prodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/prodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/prodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/prodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/prodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/prodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/prodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/prodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/prodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/prodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/prodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/prodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/prodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/prodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/prodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/prodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/prodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/prodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/prodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/prodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/prodata/1/pna/US6041_COMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2853	100.0	2853	36	US-09-965-631-3
2	2853	100.0	3446	36	US-09-965-631-7
3	2700	94.6	2867	29	US-09-741-151-1
4	2695	94.5	2853	38	US-10-009-332-2
5	2196	77.0	2930	65	US-60-216-821-9
6	1491	52.3	2469	41	US-10-163-316-3
7	1491	52.3	2469	73	US-60-297-863-3
8	1491	52.3	2940	41	US-10-163-316-1
9	1491	52.3	2940	73	US-60-297-863-1
10	1261	44.2	2297	39	US-10-093-463-27
11	1261	44.2	2297	39	US-10-093-463-29
12	1261	44.2	3400	65	US-60-212-658-762
13	1261	44.2	3759	68	US-60-242-679-1614
14	1091	38.2	1104	36	US-09-965-631-5
15	958	33.6	3625	67	US-09-965-631-1
16	958	33.6	27392	64	US-60-206-028-117
17	958	33.6	27392	64	US-60-206-028-117
18	958	33.6	27760	68	US-60-242-679-304
19	958	33.6	28824	29	US-09-741-151-3
20	958	33.6	31624	64	US-60-207-315-28
21	958	33.6	31624	64	US-60-207-315-28

```

c 22 958 33.6 31627 65 US-60-212-656-232 Sequence 232, App
c 23 958 33.6 32768 64 US-60-207-315-29 Sequence 29, App
c 24 958 33.6 32768 64 US-60-208-020-14 Sequence 14, App
c 25 958 33.6 32768 64 US-60-209-043-17 Sequence 17, App
c 26 958 33.6 32768 64 US-60-209-043-18 Sequence 18, App
c 27 958 33.6 32768 64 US-60-209-043-19 Sequence 19, App
c 28 958 33.6 32768 65 US-60-212-656-233 Sequence 233, App
c 29 958 33.6 32768 67 US-60-230-435-159 Sequence 159, App
c 30 958 33.6 32768 67 US-60-230-435-159 Sequence 159, App
c 31 844 29.6 1829 22 US-09-575-003-17 Sequence 17, App
c 32 844 29.6 1829 23 US-09-609-059-2 Sequence 2, App
c 33 844 29.6 1829 36 US-09-975-545-17 Sequence 17, App
c 34 844 29.6 1829 39 US-10-071-241-2 Sequence 2, App
c 35 838 29.4 31959 64 US-60-200-382-27 Sequence 27, App
c 36 838 29.4 31959 64 US-60-200-382-28 Sequence 28, App
c 37 724 25.4 1875 64 US-60-207-213-37 Sequence 37, App
c 38 724 25.4 1875 65 US-60-213-845-8 Sequence 8, App
c 39 639 22.4 741 65 US-60-213-845-538 Sequence 538, App
c 40 611 21.4 662 62 US-60-188-162-89 Sequence 1688, App
c 41 611 21.4 662 62 US-60-188-162-89 Sequence 1688, App
c 42 575 19.5 633 63 US-60-196-178-3578 Sequence 493, App
c 43 555 19.1 555 39 US-10-093-463-33 Sequence 35, App
c 44 544 19.1 555 39 US-10-093-463-33 Sequence 35, App
c 45 530 18.6 530 61 US-60-173-469-149 Sequence 149, App
```

ALIGNMENTS

```

RESULT 1
; Sequence 3, Application US/09965631
; GENERAL INFORMATION:
; APPLICANT: Frittdle, Carl Johan
; APPLICANT: Hildun, Erin
; FILE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
; TITLE REFERENCE: Lex-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-3
```

```

Query Match 100.0%; Score 2853; DB 36; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ATGCTTCTGCTGGGAGATCTTAACCTTGGCTTGGCGGAGAACCGCTGAGAGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGAGATCTTAACCTTGGCTTGGCGGAGAACCGCTGAGAGCTCTGAG 60
QY 61 CCAGAGCGGAGAGTAGTCTGTTCCCATCCGACTGGAGCCCGGAGCATTAAGCGCGCGCTAC 120
DB 61 CCAGAGCGGAGAGTAGTCTGTTCCCATCCGACTGGAGCCCGGAGCATTAAGCGCGCGCTAC 120
QY 121 TACTGGGGGGTCCCGAGAGATCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT 180
DB 121 TACTGGGGGGTCCCGAGAGATCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT 180
QY 181 CAGAGAGCATTTTACCTACACCTACGCGCGGATCTCAAGTTCTTGGCTCCCGCTTCTCC 240
DB 181 CAGAGAGCATTTTACCTACACCTACGCGCGGATCTCAAGTTCTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCGCTCCAGAGGGGCTACCGGGGGCTTTAGACCTGCGACGC 300
DB 241 ACTGAGCATCTGGGCGTCCCGCTCCAGAGGGGCTACCGGGGGCTTTAGACCTGCGACGC 300
```

```

QY 301 TCGTCTATCTGGGAGATGAGAGCCGAGCGGAGCTGCTGCTGAGAGCTGTGC 360
DB 301 TCGTCTATCTGGGAGATGAGAGCCGAGCGGAGCTGCTGCTGAGAGCTGTGC 360
QY 361 GGGGGGCTCCCGGAGAGCTTTGGCTACCGAGAGCGCGGAGTATGATTTAGCCCTGCCC 420
DB 361 GGGGGGCTCCCGGAGAGCTTTGGCTACCGAGAGCGCGGAGTATGATTTAGCCCTGCCC 420
QY 421 AATGCTACCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 480
DB 421 AATGCTACCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 480
QY 481 GGTGTTCCGGGAGCGGCTTCCGAGAGCCCACTCTGCTCGGAGTGGCTCGGCTG 540
DB 481 GGTGTTCCGGGAGCGGCTTCCGAGAGCCCACTCTGCTCGGAGTGGCTCGGCTG 540
QY 541 AACCCTGCATCTACGAGGCGCTTGGAGCCCTTACAAAGCGCGGCGGAGGCTTGGGCGAG 600
DB 541 AACCCTGCATCTACGAGGCGCTTGGAGCCCTTACAAAGCGCGGCGGAGGCTTGGGCGAG 600
QY 601 AGTGTAGCGCGGAGCGGAGGCTTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 660
DB 601 AGTGTAGCGCGGAGCGGAGGCTTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 660
QY 661 GAGAGCGTGTGTGTGCGGAGAGTCAATGTGTCAAGTTCAAGCGCGGAGCGGAGCGGAGCAT 720
DB 661 GAGAGCGTGTGTGTGCGGAGAGTCAATGTGTCAAGTTCAAGCGCGGAGCGGAGCGGAGCAT 720
QY 721 TATCTGTGAGCGCTGCTGCGAAAGCGGCGGAGCTTACCGCATCCAGATCTCTAAC 780
DB 721 TATCTGTGAGCGCTGCTGCGAAAGCGGCGGAGCTTACCGCATCCAGATCTCTAAC 780
QY 781 CCCATCAACATGTTGTGTGTCGAAGGTGCTTCTTAAAGATCGTGAATCCGGGCCCAAG 840
DB 781 CCCATCAACATGTTGTGTGTCGAAGGTGCTTCTTAAAGATCGTGAATCCGGGCCCAAG 840
QY 841 GTACCGGCAATGCGGCGCTGCGAAAGCGGCGGAGCTTGTGCTGCGGAGAGAGCTGAC 900
DB 841 GTACCGGCAATGCGGCGCTGCGAAAGCGGCGGAGCTTGTGCTGCGGAGAGAGCTGAC 900
QY 901 AAGTGTAGTGAACAGCACCGGAGTGTGAGGAGTGTGATGCTGCTTACAGGAGAGAG 960
DB 901 AAGTGTAGTGAACAGCACCGGAGTGTGAGGAGTGTGATGCTGCTTACAGGAGAGAG 960
QY 961 CTGTGTGAGCGGAGCGGAGCTGTGACACCGTGGGAGTGTGATGCTGCTTACAGGAGAG 1020
DB 961 CTGTGTGAGCGGAGCGGAGCTGTGACACCGTGGGAGTGTGATGCTGCTTACAGGAGAG 1020
QY 1021 CCCAAGAGAGCTGCTGTGTCTATTGAGAGCATGGGCTTCCATCAGCTTACAGAGCTG 1080
DB 1021 CCCAAGAGAGCTGCTGTGTCTATTGAGAGCATGGGCTTCCATCAGCTTACAGAGCTG 1080
QY 1081 CAGAGAGTGGGAGCGGAGTGTGATGAGAGCGGAGTGTGAGAGTGTGAGAGAGTGTG 1140
DB 1081 CAGAGAGTGGGAGCGGAGTGTGATGAGAGCGGAGTGTGAGAGTGTGAGAGAGTGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCAACCAATGATGTCCCGAGCTTATCCAGATGAGAGCTG 1200
DB 1141 TTTGGGAAGCTCCGAGCAACCAATGATGTCCCGAGCTTATCCAGATGAGAGCTG 1200
QY 1201 AACCCCTGTGAGCTGTGAGTGTGATGAGAGCGGAGTGTGAGAGTGTGAGAGAGTGTG 1260
DB 1201 AACCCCTGTGAGCTGTGAGTGTGATGAGAGCGGAGTGTGAGAGTGTGAGAGAGTGTG 1260
QY 1261 GACTGCTCTGTGAGCAACCAAGCGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1320
DB 1261 GACTGCTCTGTGAGCAACCAAGCGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1320
QY 1321 AGCTACACCTGAGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1380
DB 1321 AGCTACACCTGAGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1380
QY 1381 TACATGCACTACTGACCAAGCTGTGTGTGACCGGAGGAGGCAAGAGTGTGTG 1440
DB 1381 TACATGCACTACTGACCAAGCTGTGTGTGACCGGAGGAGGCAAGAGTGTGTG 1440
```

```
|||||
Db 1381 TACATGAGTACTGACACCAAGCTGTGTGACCGGGAAGGCCAAAGGACACATGTGTGC 1440
QY 1441 CAGACCCGACCTTCCCTGGGCGATGAGCAGCGTGTGGAGGGCAAGCTCTCC 1500
Db 1441 CAGACCCGACCTTCCCTGGGCGATGAGCAGCGTGTGGAGGGCAAGCTCTCC 1500
QY 1501 AAAGGGGCGTGGGTGGAGAGACAAACCTTCAACAGCAGAGGTGTATGTTCTCTGGCC 1560
Db 1501 AAAGGGGCGTGGGTGGAGAGACAAACCTTCAACAGCAGAGGTGTATGTTCTCTGGCC 1560
QY 1561 AAATGGGATCCCTATGGGCGGCTGTGGCGACATGTGTGGGGGCTGTGCAGCTGTCCAGG 1620
Db 1561 AAATGGGATCCCTATGGGCGGCTGTGGCGACATGTGTGGGGGCTGTGCAGCTGTCCAGG 1620
QY 1621 AGGCAGTGACCAACCCCAACCCCTGTGCAAGGAGGAGTGTGAGGAGAGTGAAGGAG 1680
Db 1621 AGGCAGTGACCAACCCCAACCCCTGTGCAAGGAGGAGTGTGAGGAGAGTGAAGGAG 1680
QY 1681 AAATACCGATCTTGCATCTGTGAAGCCCTGCCCCAGCTCAGCTCCGAAAGAGCTTCCGG 1740
Db 1681 AAATACCGATCTTGCATCTGTGAAGCCCTGCCCCAGCTCAGCTCCGAAAGAGCTTCCGG 1740
QY 1741 GAGGACAGTGTAGGCTTTCAAGGCTTCAACAGCAGACAGCAGCAGCAGCTCACTCTCGCC 1800
Db 1741 GAGGACAGTGTAGGCTTTCAAGGCTTCAACAGCAGACAGCAGCAGCAGCTCACTCTCGCC 1800
QY 1801 GTGGCATGGGTGCCAAGTACTCCGGCGTGTCCCGGGAGCAAGTGAAGCTCATCTGC 1860
Db 1801 GTGGCATGGGTGCCAAGTACTCCGGCGTGTCCCGGGAGCAAGTGAAGCTCATCTGC 1860
QY 1861 CGAGGCAATGGGACGCTGCTACTTCTATGTGTGTGACACCCAAAGGTGTGAGCGCAGCTG 1920
Db 1861 CGAGGCAATGGGACGCTGCTACTTCTATGTGTGTGACACCCAAAGGTGTGAGCGCAGCTG 1920
QY 1921 TGTCTCTGACTCTCACTCCGCTGTGTGTGCAAGGCAATGATCAAGGCTGGCTGTAT 1980
Db 1921 TGTCTCTGACTCTCACTCCGCTGTGTGTGCAAGGCAATGATCAAGGCTGGCTGTAT 1980
QY 1981 GGGAACTGGGCTCCCAAGAGAGATTCGACAGTGTGGGGTGTGTGGGGGAGACAATAG 2040
Db 1981 GGGAACTGGGCTCCCAAGAGAGATTCGACAGTGTGGGGTGTGTGGGGGAGACAATAG 2040
QY 2041 ACCTGCAAGAAGGTGACTGACTCTTCAACAGCCATCAATGCTTCAATTTCTGTGTG 2100
Db 2041 ACCTGCAAGAAGGTGACTGACTCTTCAACAGCCATCAATGCTTCAATTTCTGTGTG 2100
QY 2101 GCCATCCCGCAGGGCGCTCAAGATCGACATCCGCAAGCGGTGTACAAAGGGCTGATC 2160
Db 2101 GCCATCCCGCAGGGCGCTCAAGATCGACATCCGCAAGCGGTGTACAAAGGGCTGATC 2160
QY 2161 GGGGATGACACTACTGCTGTGAAGAACAGCAGGCAAGTACTGCTCAACGGGCAAT 2220
Db 2161 GGGGATGACACTACTGCTGTGAAGAACAGCAGGCAAGTACTGCTCAACGGGCAAT 2220
QY 2221 TTCGTGTGTGCGGGGTGGAGGCGGACCTGTTGAAGGAGAGTGTGTGAGGAGTGTG 2280
Db 2221 TTCGTGTGTGCGGGGTGGAGGCGGACCTGTTGAAGGAGAGTGTGTGAGGAGTGTG 2280
QY 2281 GGCACGGGACAGCGGTGAGAGCTGACAGGCTTCCGGCCATCTGTGACCGGTGACC 2340
Db 2281 GGCACGGGACAGCGGTGAGAGCTGACAGGCTTCCGGCCATCTGTGACCGGTGACC 2340
QY 2341 GTGGAGGTCTCTTCCGTGGGGAAGTGAACCGCCCGGGTCCGTACTCTTCTATCTG 2400
Db 2341 GTGGAGGTCTCTTCCGTGGGGAAGTGAACCGCCCGGGTCCGTACTCTTCTATCTG 2400
QY 2401 CCAAGAGAGCTCTGGGAGAGCAAGTCTCTATCCCAAGGAGCCCGGGGAGCCCTCTGTC 2460
Db 2401 CCAAGAGAGCTCTGGGAGAGCAAGTCTCTATCCCAAGGAGCCCGGGGAGCCCTCTGTC 2460
QY 2461 TTGCAACAAGGCTCTCAAGCTTCAACAGGTGAGGAGCGGAGGAGGAGGAGGAGGAGG 2520
|||||
```

```
Db 2461 TTGCAACAAGGCTCTCAAGCTTCAACAGGTGAGGAGCGGAGGAGGAGGAGGAGGAGG 2520
QY 2521 GCACGCTGGGTGGTGGAGAGCTGGGGGCGGTGCTCCGAGCTGGGAGTGGGCTGGAG 2580
Db 2521 GCACGCTGGGTGGTGGAGAGCTGGGGGCGGTGCTCCGAGCTGGGAGTGGGAGTGGAG 2580
QY 2581 AAGCGGGGCGTGGAGTGTGCGGGGCTCCGGCGGAGCGGAGCGGTGCTGATGATGA 2640
Db 2581 AAGCGGGGCGTGGAGTGTGCGGGGCTCCGGCGGAGCGGAGCGGTGCTGATGATGA 2640
QY 2641 GCCATCGGCGCGTGGAGACAAACCTGCGGGAGAGCCCTGACCTGGAGAGCTGAGC 2700
Db 2641 GCCATCGGCGCGTGGAGACAAACCTGCGGGAGAGCCCTGACCTGGAGAGCTGAGC 2700
QY 2701 GCTGTGATCCTGTCTTCAAGAGCTGCGCGGGGATTTCAAGAGGCTCACTCAAGTGT 2760
Db 2701 GCTGTGATCCTGTCTTCAAGAGCTGCGCGGGGATTTCAAGAGGCTCACTCAAGTGT 2760
QY 2761 GTGGGCAAGGAGCGGCGCTGTGGCGCGGAGCAGTGCATCTTGCACCGGAGCCAG 2820
Db 2761 GTGGGCAAGGAGCGGCGCTGTGGCGCGGAGCAGTGCATCTTGCACCGGAGCCAG 2820
QY 2821 GAGCTGACCTTCTGCTCTGAGGCGGTGCTGA 2853
Db 2821 GAGCTGACCTTCTGCTCTGAGGCGGTGCTGA 2853
```

RESULT 2

US-09-965-631-7
Sequence 7, Application US/0995631

GENERAL INFORMATION:

APPLICANT: Friddle, Carl Johan

APPLICANT: Hilbun, Erin

TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same

FILE REFERENCE: LEX-0241-USA

CURRENT FILING DATE: 2001-09-27

PRIOR APPLICATION NUMBER: US 60/236,689

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 3446

TYPE: DNA

ORGANISM: homo sapiens

US-09-965-631-7

Query Match 100.0%; Score 2853; DB 36; Length 3446;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCTTCTGCTGGGCACTCAACCTGCTTGTGCGCGGGGAACCGCTGAGGCTCTGAG 60
Db 397 ATGCTTCTGCTGGGCACTCAACCTGCTTGTGCGCGGGGAACCGCTGAGGCTCTGAG 456
QY 61 CCAGAGCGGGAGGTAGTGTTCCTCATCCGACTGAGCCCGGACATTAACGCGCGCTAC 120
Db 457 CCAGAGCGGGAGGTAGTGTTCCTCATCCGACTGAGCCCGGACATTAACGCGCGCTAC 516
QY 121 TACTGGGGGGTCCCGAGGAGTCCGGGAGTCAAGGAGTCAATTTTCAGATCAGAGCATTT 180
Db 517 TACTGGGGGGTCCCGAGGAGTCCGGGAGTCAAGGAGTCAATTTTCAGATCAGAGCATTT 576
QY 181 CAGAGAGACTTTTACCTACCTGACGCGGAGTGTGCTTGTGCTCCCGCTTCTCC 240
Db 577 CAGAGAGACTTTTACCTACCTGACGCGGAGTGTGCTTGTGCTCCCGCTTCTCC 636
QY 241 ACTGAGCATGTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGAGCTGCGAGCG 300
Db 637 ACTGAGCATGTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGAGCTGCGAGCG 696
QY 301 TGTCTCATTTCTGGGACGTGAAGCGCGGAGCTGCTGCTGTGAGCTGTGC 360
|||||
```

Db 697 TGGTTTATTCTGGGAGCTGAACCGGAGCGGACTGTTGGCTGCTGAGCCCTGCG 756
OY 361 GGGGGGCTCCGGGAGACCTTTTGGCTACCGAGGCGCCGATATGTATTAGCCCGCTGCC 420
Db 757 GGGGGGCTCCGGGAGACCTTTTGGCTACCGAGGCGCCGATATGTATTAGCCCGCTGCC 816
OY 421 AATGCTAAGCGCGCGCGCGAGCGGACACAGCCAGGCGGACACCTTCTCCAGGCGCG 480
Db 817 AATGCTAAGCGCGCGCGCGAGCGGACACAGCCAGGCGGACACCTTCTCCAGGCGCG 876
OY 481 GGTGTTCCGGGCGGCGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCCTCGGGGCTGG 540
Db 877 GGTGTTCCGGGCGGCGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCCTCGGGGCTGG 936
OY 541 AACCCTGCTCTACGCGGCGCTGAGACCTTACAGCCGCGGCGGCGGCTTTGGGGAG 600
Db 937 AACCCGCGCATCTTACGCGGCGCTGAGACCTTACAGCCGCGGCGGCGGCTTTGGGGAG 996
OY 601 AGTCGAGCGGCGGAGGCTGAGGCGCGCAAGCTTTGCTGCTATCCGCGGTACGTG 660
Db 997 AGTCGAGCGGCGGAGGCTTGGGCGCGCAAGCTTTGCTGCTATCCGCGGTACGTG 1056
OY 661 GAGACGCTGCTGCTGCGGAGCATATGCTCAAGTTCACGCGCGGACCTGAGACAT 720
Db 1057 GAGACGCTGCTGCTGCGGAGCATATGCTCAAGTTCACGCGCGGACCTGAGACAT 1116
OY 721 TATCTGCTGAGCTGCTGCGAAGCGGGGCGGACCTACCGGCAATCCGAGATCTCAAC 780
Db 1117 TATCTGCTGAGCTGCTGCGAAGCGGGGCGGACCTACCGGCAATCCGAGATCTCAAC 1176
OY 781 CCCATCAACATGTTGTGTCAAGGTGCTGCTTCTTAGAGATGCTGACTCGGGGCCAAG 840
Db 1177 CCCATCAACATGTTGTGTCAAGGTGCTGCTTCTTAGAGATGCTGACTCGGGGCCAAG 1236
OY 841 GTCAACGCGCAATGCGGCGCTGAGCGCTGCGCAACTTCTGTGCTGCGGAGAGAGCTGAGC 900
Db 1237 GTCAACGCGCAATGCGGCGCTGAGCGCTGCGCAACTTCTGTGCTGCGGAGAGAGCTGAGC 1296
OY 901 AAGTGTGTGACAAAGACCCCGAGTACTGGGACACTGCGCATCTCTCTTCAACAGGACGAGC 960
Db 1297 AAGTGTGTGACAAAGACCCCGAGTACTGGGACACTGCGCATCTCTCTTCAACAGGACGAGC 1356
OY 961 CTGTGTGAGCCACACCTGTGACACCTTGGGCAATGGCTGATGGGTACATGTGTGAC 1020
Db 1357 CTGTGTGAGCCACACCTGTGACACCTTGGGCAATGGCTGATGGGTACATGTGTGAC 1416
OY 1021 CCCAAGGAGGCTGCTGTGCTTGAAGAGATGGGCTTCATCAGGCTTCAACACCTGCC 1080
Db 1417 CCCAAGGAGGCTGCTGTGCTTGAAGAGATGGGCTTCATCAGGCTTCAACACCTGCC 1476
OY 1081 CACGAGCTGGGCGCAGTGTCAACATGCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1140
Db 1477 CACGAGCTGGGCGCAGTGTCAACATGCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1536
OY 1141 TTTGGGAGGCTCGAGCGCAACACATATGTCCCCGACCTCATCCAGATGACGCTGCC 1200
Db 1537 TTTGGGAGGCTCGAGCGCAACACATATGTCCCCGACCTCATCCAGATGACGCTGCC 1596
OY 1201 AACCCTGCTGAGCTGAGTGTGCTGATCATCAGACTTCCGAGACGCGGACGAGT 1260
Db 1597 AACCCTGCTGAGCTGAGTGTGCTGATCATCAGACTTCCGAGACGCGGACGAGT 1556
OY 1261 GACTGCTCTCTGAGCAACCCAGCAAGCCCATCTCCCTGCCGAGATCTGCGGGGCGC 1320
Db 1657 GACTGCTCTCTGAGCAACCCAGCAAGCCCATCTCCCTGCCGAGATCTGCGGGGCGC 1716
OY 1321 AGCTACACCTGAGCCAGCAAGTGTGAGTGTGCTTTGGCGTGGCTCCAAAGCCCTGTCT 1380
Db 1717 AGCTACACCTGAGCCAGCAAGTGTGAGTGTGCTTTGGCGTGGCTCCAAAGCCCTGTCT 1776
OY 1381 TACATGACGATGCTGACCAACGCTGTGGTGACCGGGAAGGCCAAGGAGACAGATGTGTGC 1440
Db 1777 TACATGACGATGCTGACCAACGCTGTGGTGACCGGGAAGGCCAAGGAGACAGATGTGTGC 1836

OY 1441 CAGACCCGCGCATCTCCCTGGGGCGATGGACCAAGCTGTGGCGAGGCGAGCTCTGCTTC 1500
Db 1837 CAGACCCGCGCATCTCCCTGGGGCGATGGACCAAGCTGTGGCGAGGCGAGCTCTGCTTC 1896
OY 1501 AAGGCGGCTGCTGAGAGAGACACAACTCAACAGCAGAGGTGATGCTTCTGGGCC 1560
Db 1897 AAGGCGGCTGCTGAGAGAGACACAACTCAACAGCAGAGGTGATGCTTCTGGGCC 1956
OY 1561 AATGAGATCTCTATGCGCCCTGCTGCGCAATGTGTGGGGCGGCTGAGCTGGCCAGG 1620
Db 1957 AATGAGATCTCTATGCGCCCTGCTGCGCAATGTGTGGGGCGGCTGAGCTGGCCAGG 2016
OY 1621 AGGCACTGACCAACCCACCTCTGCGCAAGCGGGGCGAATGCTGCGAGGAGTGTGAGGTTG 1680
Db 2017 AGGCACTGACCAACCCACCTCTGCGCAAGCGGGGCGAATGCTGCGAGGAGTGTGAGGTTG 2076
OY 1681 AATATCGGATCTCTCAATCTGAGGCGCTGCCAGCTGACGCTTCCGGAAGAGCTTCCGG 1740
Db 2077 AATATCGGATCTCTCAATCTGAGGCGCTGCCAGCTGACGCTTCCGGAAGAGCTTCCGG 2136
OY 1741 GAGGAGCGTGTGAGGCTTTCAGGGGTACAAACAGCAGCAACCGGCTCACTCTGCC 1800
Db 2137 GAGGAGCGTGTGAGGCTTTCAGGGGTACAAACAGCAGCAACCGGCTCACTCTGCC 2196
OY 1801 GTGGCATGGTGGCCCAAGTACTCCGGGCTGTCTCCCGGGGACAGTGCAGCTCATCTGC 1860
Db 2197 GTGGCATGGTGGCCCAAGTACTCCGGGCTGTCTCCCGGGGACAGTGCAGCTCATCTGC 2256
OY 1861 CGAGCAATGCTGCTGCTTCTATGTGTGCGACCCCAAGGTTGGGAGCGGAGCTGC 1920
Db 2257 CGAGCAATGCTGCTGCTTCTATGTGTGCGACCCCAAGGTTGGGAGCGGAGCTGC 2316
OY 1921 TGCCTCTGACTCCACTCCGCTGCTGTGTGCCAAGGCAAGTGCATCAAGGCTGCTGAT 1980
Db 2317 TGCCTCTGACTCCACTCCGCTGCTGTGTGCCAAGGCAAGTGCATCAAGGCTGCTGAT 2376
OY 1981 GGGAACTGGGCTCCAGAGAGATTTGACAAAGTGTGGGGTGTGGGGGAGCAATAG 2040
Db 2377 GGGAACTGGGCTCCAGAGAGATTTGACAAAGTGTGGGGTGTGGGGGAGCAATAG 2436
OY 2041 ACCTGCAAGAGGCTGCTGCTTTCACCAAGCCCATGCTATGCTCAATTTCTGCTGAT 2100
Db 2437 ACCTGCAAGAGGCTGCTGCTTTCACCAAGCCCATGCTATGCTCAATTTCTGCTGAT 2496
OY 2101 GCCATCCCGCAGGCGCTCAAGCATCGGACGCGCGGTTCACAAAGGCTGATC 2160
Db 2497 GCCATCCCGCAGGCGCTCAAGCATCGGACGCGCGGTTCACAAAGGCTGATC 2556
OY 2161 GGGGATGACACTACTGCTGCTGTGAAGAACGCGCAAGGCAAGTACCTCTCAACGGGAT 2220
Db 2557 GGGGATGACACTACTGCTGCTGTGAAGAACGCGCAAGGCAAGTACCTCTCAACGGGAT 2616
OY 2221 TTCGTGTGTGGGCGGTGAGGCGGAGCTGTGTGAAGGCGGAGTCTCGGTACAGC 2280
Db 2617 TTCGTGTGTGGGCGGTGAGGCGGAGCTGTGTGAAGGCGGAGTCTCGGTACAGC 2676
OY 2281 GGCAGCGGACAGCGGTGAGAGCTGCAAGCTTCCCGGCCATCTCTGAGACCGCTGACC 2340
Db 2677 GGCAGCGGACAGCGGTGAGAGCTGCAAGCTTCCCGGCCATCTCTGAGACCGCTGACC 2736
OY 2341 GTGGAGTCTCTCTGCTGGGGAAGATGACACCGCGCGGATCTGCTCTATGTG 2400
Db 2737 GTGGAGTCTCTCTGCTGGGGAAGATGACACCGCGCGGATCTGCTCTATGTG 2796
OY 2401 CCCAAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGAGACCCCGGGGAGCCCTGTCTC 2460
Db 2797 CCCAAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGAGACCCCGGGGAGCCCTGTCTC 2856
OY 2461 TTGCACACAGCGCTCTGAGCTTCAACAGGTGTGAGAGCGGAGCAGAGCGCCCT 2520
Db 2857 TTGCACACAGCGCTCTGAGCTTCAACAGGTGTGAGAGCGGAGCAGAGCGCCCT 2916

Db 1444 CAGACCGGACCTTCCCTGGGCGGATGGACAGCTGT 36GAGGCAAGCTTGCCCTC 1503
Qy 1501 AAGGGGCGCTGGCTGGAGAGACAACTTCAACACAGAGGTGATGGTTCCTGGGCG 1560
Db 1504 AAGGGGCGCTGGCTGGAGAGACAACTTCAACACAGAGGTGATGGTTCCTGGGCG 1563
Qy 1561 AATGGGATCCCTATGAGGCGCTGGCGACATGTGTGGGGGCTGACAGTGGCGAG 1620
Db 1564 AATGGGATCCCTATGAGGCGCTGGCGACATGTGTGGGGGCTGACAGTGGCGAG 1623
Qy 1621 AGGCAAGTGCACCAACCCACCTCCCAACGGGGGCAAGTACTGGAGGAGTGAAGGTG 1680
Db 1624 AGGCAAGTGCACCAACCCACCTCCCAACGGGGGCAAGTACTGGAGGAGTGAAGGTG 1683
Qy 1681 AATATCCGATCTGCAATCTGGAGGCGCTGGCGGCAAGCTCCGGAAGAGCTTCGGG 1740
Db 1684 AATATCCGATCTGCAATCTGGAGGCGCTGGCGGCAAGCTCCGGAAGAGCTTCGGG 1743
Qy 1741 GAGGAGCATGTGAGGCTTTCACAGGGCTACACACAGCAGCAACCGGCTCACTCTGGCC 1800
Db 1744 GAGGAGCATGTGAGGCTTTCACAGGGCTACACACAGCAGCAACCGGCTCACTCTGGCC 1803
Qy 1801 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTGCACAGCTCATCTGC 1860
Db 1804 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTGCACAGCTCATCTGC 1863
Qy 1861 CGAGCCATGGCACTGGCTACTTCTATGTGCTGGACCCCAAGGTGGTGGAGCGCAGCTG 1920
Db 1864 CGAGCCATGGCACTGGCTACTTCTATGTGCTGGACCCCAAGGTGGTGGAGCGCAGCTG 1923
Qy 1921 TGCTCTCTGACTGACCTCCGCTGTGTGTCCAAAGGCAAGTGCATCAAGCTGCTGTGAT 1980
Db 1924 TGCTCTCTGACTGACCTCCGCTGTGTGTCCAAAGGCAAGTGCATCAAGCTGCTGTGAT 1983
Qy 1981 GGGAGCTGGGCTCCAGAGAGATTCGACAGTGGGGTGTGGGGGAGACAAATAG 2040
Db 1984 GGGAGCTGGGCTCCAGAGAGATTCGACAGTGGGGTGTGGGGGAGACAAATAG 2043
Qy 2041 AGCTGCAAGAGGTGACTGAGCTTTCACCAAGGCCATGCTGCTCAATTTTGTGTG 2100
Db 2044 AGCTGCAAGAGGTGACTGAGCTTTCACCAAGGCCATGCTGCTCAATTTTGTGTG 2103
Qy 2101 GCCATCCCCGAGCGGCTCAAGCATGACATCGCCAGCGGTTTACAAAGGCTGTATC 2160
Db 2104 GCCATCCCCGAGCGGCTCAAGCATGACATCGCCAGCGGTTTACAAAGGCTGTATC 2163
Qy 2161 GGGGATGCAACTGACTGGCTCTGAAGACAGCAAGCAAGTACTGCTCAACGGGCAAT 2220
Db 2164 GGGGATGCAACTGACTGGCTCTGAAGACAGCAAGCAAGTACTGCTCAACGGGCAAT 2223
Qy 2221 TTGCTGTGTGCGGCGGTGAGACGGGACCTGTGTGTGAAGGCAATGCTGTGCGGTACAGC 2280
Db 2224 TTGCTGTGTGCGGCGGTGAGACGGGACCTGTGTGTGAAGGCAATGCTGTGCGGTACAGC 2283
Qy 2281 GGCACGGGCAAGGGGTGGAGAGCTTCCGGGCGCATTCCTGGAGCGCTGAGC 2340
Db 2284 GGCACGGGCAAGGGGTGGAGAGCTTCCGGGCGCATTCCTGGAGCGCTGAGC 2343
Qy 2341 GTGAGGCTCTCTCCGTGGGAGATGACACCGCGGGGTCTGCTACTCTTTATCTG 2400
Db 2344 GTGAGGCTCTCTCCGTGGGAGATGACACCGCGGGGTCTGCTACTCTTTATCTG 2403
Qy 2401 CCCAAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGAGACCCCGGGGACCTCTGTG 2460
Db 2404 CCCAAAGAGCTCGGAGAGCAAGTCTCTATCCCAAGAGACCCCGGGGACCTCTGTG 2463
Qy 2461 TTGCAACAACAGCTCTTCAAGCTTCAACAGGTGAGAGCGCGGAGACAGGCGCCCT 2520
Db 2464 TTGCAACAACAGCTCTTCAAGCTTCAACAGGTGAGAGCGCGGAGACAGGCGCCCT 2523
Qy 2521 GCACGCTGGGTGGAGCTGGAGCTGGTGTCTCCGAGAGCTCGGCAAGTGGCTGAG 2580
Db 2524 GCACGCTGGGTGGAGCTGGAGCTGGTGTCTCCGAGAGCTGGGCAAGTGGCTGAG 2583

Qy 2581 AAGCGGCGGTGAGACTGTGCGGGCTCCGCGGAGCGCAAGTCCCTGCTGTATGCA 2640
Db 2584 AAGCGGCGGTGAGACTGTGCGGGGCTCCGCGGAGCGCAAGGTCCTGCTGTATGCA 2643
Qy 2641 GCCCATGGGCGGTGAGACACAAAGCCTGCGGGGAGGCTTCCCTACCTGGAGACTAGC 2700
Db 2644 GCCCATGGGCGGTGAGACACAAAGCCTGCGGGGAGGCTTCCCTACCTGGAGACTAGC 2703
Qy 2701 GCTGTGACCTGCTCCAAAGCTGGGCGGCGGAGATTTCAGAGGCTCACTCAATGT 2760
Db 2704 GCTGTGACCTGCTCCAAAGCTGGGCGGCGGAGATTTCAGAGGCTCACTCAATGT 2763
Qy 2761 GTGGGCAAGAGGCGGCTGCTGGCCCGGAGCAAGTGCACATTTGACACCGCAAGCC 2820
Db 2764 GTGGGCAAGAGGCGGCTGCTGGCCCGGAGCAAGTGCACATTTGACACCGCAAGCC 2823
Qy 2821 GAGCTGAGCTTCTGCTCTTCAAGGCGCTGTGCA 2853
Db 2824 GAGCTGAGCTTCTGCTCTTCAAGGCGCTGTGCA 2856

RESULT 4

US-10-009-332-2
; Sequence 2, Application US/10009332

; GENERAL INFORMATION:

; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.

; APPLICANT: Kazusa DNA Research Institute

; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY

; FILE REFERENCE: 067541

; CURRENT FILING DATE: 2001-12-10

; PRIOR FILING DATE: 1999-11-11

; PRIOR APPLICATION NUMBER: JPA Hei 11-321740

; PRIOR FILING DATE: 2000-05-16

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2853

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-009-332-2

Query Match 94.5%; Score 2695; DB 38; Length 2853;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 57 TGAGCGAGAGCGGAGAGTGTGCTTCCCATCCGACTGGACCGGACATTAACGGCGCG 116
Db 57 TGAGCGAGAGCGGAGAGTGTGCTTCCCATCCGACTGGACCGGACATTAACGGCGCG 116
Qy 117 CTACTACTGCGGGGTCCCGAGGACTCCGGGATGAGGACTCATTTTCAGATACAGC 176
Db 117 CTACTACTGCGGGGTCCCGAGGACTCCGGGATGAGGACTCATTTTCAGATACAGC 176
Qy 117 ATTTCAGAGGACTTTTACTTACACTGACGGCGGATGCTAGTGTGGCTCCGCCCTT 236
Db 117 ATTTCAGAGGACTTTTACTTACACTGACGGCGGATGCTAGTGTGGCTCCGCCCTT 236
Qy 237 CTCACCTGAGCACTGTGGGCTCCCTCCAGGGGCTCAACGGGGGCTTTAGACCTGCG 296
Db 237 CTCACCTGAGCACTGTGGGCTCCCTCCAGGGGCTCAACGGGGGCTTTAGACCTGCG 296
Qy 297 ACCTGCTCTTCTATTTCTGGGAGCTGAACGCGGAGCGGACTGTTCGCTGTGAGCT 356
Db 297 ACCTGCTCTTCTATTTCTGGGAGCTGAACGCGGAGCGGACTGTTCGCTGTGAGCT 356
Qy 357 GTGCGGGGGGCTCCGGAGGCTTTGGCTACAGAGGCGCGGATGTGATTAAGCCGCT 416
Db 357 GTGCGGGGGGCTCCGGAGGCTTTGGCTACAGAGGCGCGGATGTGATTAAGCCGCT 416
Qy 417 GCCCAATGTCTAGGCGCGCGGCGGAGCGCAACAGCCAGGGCGGACACCTTCTCAGCG 476

|||||
Db 417 GCCCAATGCTAGCGCGCGCGCGCGCAAGCCCAAGCGCGCAACACTTCTCCAGCG 476
OY 477 CCGGGGTTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCCCTCGG 536
Db 477 CCGGGGTTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGGCCCTCGG 536
OY 537 CTGGAACCCCGCATCTCTACGGGCGCTTGACCTTACAAAGCCGGCGGGCGGGCTTCG 596
Db 537 CTGGAACCCCGCATCTCTACGGGCGCTTGACCTTACAAAGCCGGCGGGCGGGCTTCG 596
OY 597 GGAGAGTCGTAGCGCGCGCGCGCTGCGCGCGCAAGCGCTTGTCTATCCCGCGTA 656
Db 597 GGAGAGTCGTAGCGCGCGCGCGCTGCGCGCGCAAGCGCTTGTCTATCCCGCGTA 656
OY 657 CGAGGAGAGCTGTGGTTCGGAGCGAGTCAATGATAGTTCACGGCGCGGCACTGGA 716
Db 657 CGAGGAGAGCTGTGGTTCGGAGCGAGTCAATGATAGTTCACGGCGCGGCACTGGA 716
OY 717 ACATTAATCTGCTAGCGCTGTGCAACGGCGCGGCTTACCGCATCCAGCATCT 776
Db 717 ACATTAATCTGCTAGCGCTGTGCAACGGCGCGGCTTACCGCATCCAGCATCT 776
OY 777 CAACCCCATTAACATGCTGTGTCAAGGTGCTCTTCTTAAAGATCTGACTCCGGGCT 836
Db 777 CAACCCCATTAACATGCTGTGTCAAGGTGCTCTTCTTAAAGATCTGACTCCGGGCT 836
OY 837 CAAGGTACCGGCAATGCGGCGCTGACGCTGCGCAACTTCTGCGCTGGCGAGAAAGCT 896
Db 837 CAAGGTACCGGCAATGCGGCGCTGACGCTGCGCAACTTCTGCGCTGGCGAGAAAGCT 896
OY 897 GAACAAAGTAGTAGAAGACACCCGAGTACTGGAGACATGCCATCTCTTACACGAGA 956
Db 897 GAACAAAGTAGTAGAAGACACCCGAGTACTGGAGACATGCCATCTCTTACACGAGA 956
OY 957 GGAACCTGTGTGGAGCCACCACTGTACACCTGGGCAATGCTGATGTGGATACATGTG 1016
Db 957 GGAACCTGTGTGGAGCCACCACTGTACACCTGGGCAATGCTGATGTGGATACATGTG 1016
OY 1017 TGACCCCAAGAGAGTGTCTGTCTCATTTAGAGAGCATGGGCTTCCATGACCTTCAC 1076
Db 1017 TGACCCCAAGAGAGTGTCTGTCTCATTTAGAGAGCATGGGCTTCCATGACCTTCAC 1076
OY 1077 TGCCCAAGAGCTGGCCACAGTGTGAACATGCCCATGACCAATGTGAAGTCTGTAGA 1136
Db 1077 TGCCCAAGAGCTGGCCACAGTGTGAACATGCCCATGACCAATGTGAAGTCTGTAGA 1136
OY 1137 GGTGTTTGGGAAGCTCCGAGGCAACCAATGATGTCCCGACCTTCATCCAGATGACCG 1196
Db 1137 GGTGTTTGGGAAGCTCCGAGGCAACCAATGATGTCCCGACCTTCATCCAGATGACCG 1196
OY 1197 TGCCCAAGAGCTGTGAGCTGTGAGTGTGCAATCATCAGCACTTCTGGAGAGCGAGA 1256
Db 1197 TGCCCAAGAGCTGTGAGCTGTGAGTGTGCAATCATCAGCACTTCTGGAGAGCGAGA 1256
OY 1257 CGGTGACTGCTCTGTGAGCAACCAAGAGCCCATCTCCCTGCGGAGATCTCGCGG 1316
Db 1257 CGGTGACTGCTCTGTGAGCAACCAAGAGCCCATCTCCCTGCGGAGATCTCGCGG 1316
OY 1317 CGGCACTACACCTGAGCGAGAGTGGAGTGTGCTTTTGGCGTGGCTCCAAAGCCCTG 1376
Db 1317 CGGCACTACACCTGAGCGAGAGTGGAGTGTGCTTTTGGCGTGGCTCCAAAGCCCTG 1376
OY 1377 TCCCTACATGAGTACTGACCAAGTGTGTGACCGGAGAGGCAAGGACAGATGAT 1436
Db 1377 TCCCTACATGAGTACTGACCAAGTGTGTGACCGGAGAGGCAAGGACAGATGAT 1436
OY 1437 GTGCCAGAGCCCGCACTTCCCTGGGCGATGTGCAACAGCTGTGGCGAGGCAACTCTG 1496
Db 1437 GTGCCAGAGCCCGCACTTCCCTGGGCGATGTGCAACAGCTGTGGCGAGGCAACTCTG 1496
OY 1497 CCTCAAGGGGCTGTGGAGAGACCAACTCAACAAAGCAAGGTGATGTTCTGTG 1556
|||||

Db 1497 CCTCAAGGGGCTGTGGAGAGACACACCTCAACAAAGCAAGAGGTGATGTTCTG 1556
OY 1557 GGCCAATATGGATTCCTATGGCCCTGCTGGCGACATGTGTGGGGCTGTGACGTGGC 1616
Db 1557 GGCCAATATGGATTCCTATGGCCCTGCTGGCGACATGTGTGGGGCTGTGACGTGGC 1616
OY 1617 CAGGAGGAGTGCACCAACCCCGCTGCGCAACGGGGGCAAGTCTGCGAGGGGTAAG 1676
Db 1617 CAGGAGGAGTGCACCAACCCCGCTGCGCAACGGGGGCAAGTCTGCGAGGGGTAAG 1676
OY 1677 GGTGAATATCCGATCTGCAATCTGGAAGCCCTGCCAGCTCAGCTTCGGGAAGAGCTT 1736
Db 1677 GGTGAATATCCGATCTGCAATCTGGAAGCCCTGCCAGCTCAGCTTCGGGAAGAGCTT 1736
OY 1737 CCGGAGAGAGAGTGTGAGGCTTTTAAAGGCTACACCAAGCACACCGGCTCAGCT 1796
Db 1737 CCGGAGAGAGAGTGTGAGGCTTTTAAAGGCTACACCAAGCACACCGGCTCAGCT 1796
OY 1797 CGCGTGGCATGGTGGCCCAAGTACTCCGGGCTGTCTCCGGGGCAAGTGCAGAGTCTAT 1856
Db 1797 CGCGTGGCATGGTGGCCCAAGTACTCCGGGCTGTCTCCGGGGCAAGTGCAGAGTCTAT 1856
OY 1857 CTGCGAGCCCAATGSCACTGGCTACTTCTATGTCTGACACCAAGTGTGAGACGAC 1916
Db 1857 CTGCGAGCCCAATGSCACTGGCTACTTCTATGTCTGACACCAAGTGTGAGACGAC 1916
OY 1917 GCTGTGCTCTCTGACTCCTCAGCTCGCTGTGTGTCCAAAGGCAATGATCAAGCTGCTG 1976
Db 1917 GCTGTGCTCTCTGACTCCTCAGCTCGCTGTGTGTCCAAAGGCAATGATCAAGCTGCTG 1976
OY 1977 TGATGGGAACCTGGGCTCTCAAGAAAGATGACAAAGTGTGGGGTGTGGGGAGAGCAA 2036
Db 1977 TGATGGGAACCTGGGCTCTCAAGAAAGATGACAAAGTGTGGGGTGTGGGGAGAGCAA 2036
OY 2037 TAAGAGCTGCAAGAGTGTGACTGTGACTCTTACCAAGGCCATGATGATGATCAATTTCT 2096
Db 2037 TAAGAGCTGCAAGAGTGTGACTGTGACTCTTACCAAGGCCATGATGATGATCAATTTCT 2096
OY 2097 GGTGGGCAATCCCGAGGCGGCTCAAGATGACATCCGCCAGGGGTTTCAAAAGGCT 2156
Db 2097 GGTGGGCAATCCCGAGGCGGCTCAAGATGACATCCGCCAGGGGTTTCAAAAGGCT 2156
OY 2157 GATCGGGATGACAACTGCTGCTGTGAAGAACAGCAAGCAAGTACCTGCTCAACG 2216
Db 2157 GATCGGGATGACAACTGCTGCTGTGAAGAACAGCAAGCAAGTACCTGCTCAACG 2216
OY 2217 GCATTTCGTGTGTGCGGCTGTGAGCGGAGCTGTGTGAAGGGCAAGTCTGCTGCGTA 2276
Db 2217 GCATTTCGTGTGTGCGGCTGTGAGCGGAGCTGTGTGAAGGGCAAGTCTGCTGCGTA 2276
OY 2277 CAGCGGCAAGGAGAGGAGTGTGAGAGCTGTGCAAGGCTTCCCGGCGCAATCCGGAAGCGCT 2336
Db 2277 CAGCGGCAAGGAGAGGAGTGTGAGAGCTGTGCAAGGCTTCCCGGCGCAATCCGGAAGCGCT 2336
OY 2337 GACCGTGAAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTA 2396
Db 2337 GACCGTGAAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTA 2396
OY 2397 TCTGCGCCAAAGAGCTGTGGAGAGCAAGTCTCTCATCTCCAAAGAACCCCGGGAGCCCTC 2456
Db 2397 TCTGCGCCAAAGAGCTGTGGAGAGCAAGTCTCTCATCTCCAAAGAACCCCGGGAGCCCTC 2456
OY 2457 TGTCTTGACAAAGAGCTGCTCAGCTCTCCAAACAGAGTGGAGAGCGGAGCAAGCGCC 2516
Db 2457 TGTCTTGACAAAGAGCTGCTCAGCTCTCCAAACAGAGTGGAGAGCGGAGCAAGCGCC 2516
OY 2517 CCTTGACAGCTGTGCTGTGCACTGTGGGCGCTGTCTCCGAGAGTGCAGCAAGTGGCT 2576
Db 2517 CCTTGACAGCTGTGCTGTGCACTGTGGGCGCTGTCTCCGAGAGTGCAGCAAGTGGCT 2576
OY 2577 GCAGAAAGCGGCGGTGAGTGTGCGGGCTCCCGCGGCGAGCGTCCCTGCTCTGTGA 2636
Db 2577 GCAGAAAGCGGCGGTGAGTGTGCGGGCTCCCGCGGCGAGCGTCCCTGCTCTGTGA 2636

```
QY 2637 TGCAGCCCATCGGCGCGTGGAGACACAAGCCTGCGGGGAGCCCTGCCCCACCTGGAGCT 2696
      |||||||
Db 2637 TGCAGCCCATCGGCGCGTGGAGACACAAGCCTGCGGGGAGCCCTGCCCCACCTGGAGCT 2696
QY 2697 CAGGCGCTGTGACACCTGCTCCAGAGAGCTGCGGCGGGGATTTTCAGAGCGCTCACTCAA 2756
      |||||||
Db 2697 CAGGCGCTGTGACACCTGCTCCAGAGAGCTGCGGCGGGGATTTTCAGAGCGCTCACTCAA 2756
QY 2757 GTGTGTGGCCACGAGAGCGCGGCTGTGCGCCCGGAGACCATGTGACACCGCAAGCC 2816
      |||||||
Db 2757 GTGTGTGGCCACGAGAGCGCGGCTGTGCGCCCGGAGACCATGTGACACCGCAAGCC 2816
QY 2817 CCAGAGAGCTGAGCTTCTGCGTCTGAGAGCGCTGCTGA 2853
      |||||||
Db 2817 CCAGAGAGCTGAGCTTCTGCGTCTGAGAGCGCTGCTGA 2853

RESULT 5
US-60-216-821-9
: Sequence 9, Application US/60216821
: GENERAL INFORMATION:
: APPLICANT: Delegeane, Angelo M.
: APPLICANT: Tang, Y. Tom
: APPLICANT: Walsh, Roderick T.
: APPLICANT: Lal, Preeti
: APPLICANT: Khan, Farrah A.
: APPLICANT: Nguyen, Daniel B.
: APPLICANT: Tribouley, Catherine M.
: APPLICANT: Yue, Henry
: APPLICANT: Hatalla, April
: APPLICANT: Patterson, Chandra
: APPLICANT: Lu, Dyung Alma M.
: APPLICANT: Azimzai, Yalda
: APPLICANT: Gandhi, Ameena
: APPLICANT: Baughn, Mariah R.
: APPLICANT: Yao, Monique G.
: APPLICANT: Lee, Ernestine A.
: APPLICANT: Lu, Yan
: APPLICANT: Rankumar, Jaya
: APPLICANT: Kallick, Deborah A.
: APPLICANT: Au-Young, Janice
: APPLICANT: Xu, Yuming
: TITLE OF INVENTION: PROTEASES
: FILE REFERENCE: PI-0146 P
: CURRENT APPLICATION NUMBER: US/60/216, 821
: CURRENT FILING DATE: 2000-07-07
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PERL Program
: SEQ ID NO 9
: LENGTH: 2930
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No: 7473089CBI
US-60-216-821-9

Query Match 77.0%: Score 2196; DB 65; Length 2930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
```

```
QY 181 CAGAGAGACTTTTACCTACACCTGAGCGCGATGCTCAGTTCTTTGGCTCCCGCTTCTCC 240
      |||||||
Db 255 CAGAGAGACTTTTACCTACACCTGAGCGCGATGCTCAGTTCTTTGGCTCCCGCTTCTCC 314
QY 241 ACTGAGCATCTGCGGCTCCCTCCAGAGGCTCACCGGAGGCTCTTTCAGACCTGCGAGC 300
      |||||||
Db 315 ACTGAGCATCTGCGGCTCCCTCCAGAGGCTCACCGGAGGCTCTTTCAGACCTGCGAGC 374
QY 301 TGGTTCTATTCTGGGAGCTGAGCGCGAGCGGAGCTGTCGCTGTGAGCTGTGC 360
      |||||||
Db 375 TGGTTCTATTCTGGGAGCTGAGCGCGAGCGGAGCTGTCGCTGTGAGCTGTGC 434
QY 361 GGGGGGCTCCGCGGAGACCTTTTGGCTACCGAGGCGCGGATATGTCATTAGCCCGCTGCC 420
      |||||||
Db 435 GGGGGGCTCCGCGGAGACCTTTTGGCTACCGAGGCGCGGATATGTCATTAGCCCGCTGCC 494
QY 421 AATGCTAGCGCGCGCGCGCGAGCGCAACACCGAGGCGCAACCTTCTCCAGCGCGCG 480
      |||||||
Db 495 AATGCTAGCGCGCGCGCGCGAGCGCAACACCGAGGCGCAACCTTCTCCAGCGCGCG 554
QY 481 GGTGTTCCGGGGGCGGCTTCCGAGAGACCCACCTCTGCTGCGGGGGTGGCTCGGGCTGG 540
      |||||||
Db 555 GGTGTTCCGGGGGCGGCTTCCGAGAGACCCACCTCTGCTGCGGGGGTGGCTCGGGCTGG 614
QY 541 AACCCCGCATCTTACGCGGCGCTGAGACCTTTCAGAGCGCGGGGCGGCGCTTGGGGAG 600
      |||||||
Db 615 AACCCCGCATCTTACGCGGCGCTGAGACCTTTCAGAGCGCGGGGCGGCGCTTGGGGAG 674
QY 601 AGTCGTAGCGCGCGAGGCTGCGGCGCGGAGCGCTTTCGTATCCCGGGGTAGCTG 660
      |||||||
Db 675 AGTCGTAGCGCGCGAGGCTGCGGCGCGGAGCGCTTTCGTATCCCGGGGTAGCTG 734
QY 661 GAGAGCGCTGTTGGTGGCGGAGGAGTCAATGTCAAGTTCCAGCGCGGAGCTGAGACAT 720
      |||||||
Db 735 GAGAGCGCTGTTGGTGGCGGAGGAGTCAATGTCAAGTTCCAGCGCGGAGCTGAGACAT 794
QY 721 TATCTGCTAGCGCTGTGCGCAACGGGCGCGAGCTTACCCGATCTCTCAAC 780
      |||||||
Db 795 TATCTGCTAGCGCTGTGCGCAACGGGCGCGAGCTTACCCGATCTCTCTCAAC 854
QY 781 CCCATCAATGTTGGTGGTCAAGGTCGCTTCTTAGAGATCGTACCGCGGCGCGAG 840
      |||||||
Db 855 CCCATCAATGTTGGTGGTCAAGGTCGCTTCTTAGAGATCGTACCGCGGCGCGAG 914
QY 841 GTACCGGCAATGCGGCGCTGAGCGGTGCGCAACTTCTGTGCTGCGAGAGAGAGCTGAAC 900
      |||||||
Db 915 GTACCGGCAATGCGGCGCTGAGCGGTGCGCAACTTCTGTGCTGCGAGAGAGCTGAAC 974
QY 901 AAGTGAAGTGAACAGACCCCGAGTACTGGGAGCTGCCATCTCTTCAACAGCGAGCAG 960
      |||||||
Db 975 AAGTGAAGTGAACAGACCCCGAGTACTGGGAGCTGCCATCTCTTCAACAGCGAGCAG 1034
QY 961 CTGTGTGGAGCCACGACCGCTGACACCCCTGGGCGATGGGTGATGTGGTACATGTGTGAC 1020
      |||||||
Db 1035 CTGTGTGGAGCCACGACCGCTGACACCCCTGGGCGATGGGTGATGTGGTACATGTGTGAC 1094
QY 1021 CCCAAGAGAGCTGCTCTGTCAATGTAGAGAGTGGGCTTCCATFCACTCCACCTGCC 1080
      |||||||
Db 1095 CCCAAGAGAGCTGCTCTGTCAATGTAGAGAGTGGGCTTCCATFCACTCCACCTGCC 1154
QY 1081 CACGAGCTGGGCGAGCTGTTCACATGCCCAATGACATGTAAAGTCTGTGAGAGGTG 1140
      |||||||
Db 1155 CACGAGCTGGGCGAGCTGTTCACATGCCCAATGACATGTAAAGTCTGTGAGAGGTG 1214
QY 1141 TTTGGAGAGCTCGAGCGCAACGACATGATGTCGCCGACCTCATGACATGACAGCTGCC 1200
      |||||||
Db 1215 TTTGGAGAGCTCGAGCGCAACGACATGATGTCGCCGACCTCATGACATGACAGCTGCC 1274
QY 1201 AACCCCTGTGAGCGCTGAGCTGTGCTGACATCACCGACTTCTCTGAGACAGCGGGCACGGT 1260
      |||||||
Db 1275 AACCCCTGTGAGCGCTGAGCTGTGCTGACATCACCGACTTCTCTGAGACAGCGGGCACGGT 1334
```

QY	1281	GACGCGCTCCGGAGCCAAACCACAGAGCCCATCTCCGCGCCGAGATCTGCCGGGCGCC	132.20
Db	1335	GACTGCTCTCTGGAGCAACCCACCAAGCCCATCTCCGCGCGGAGATCTGCCGGGCGCC	139.44
QY	1321	AGCTACACCCCTGAGCCAGCAGTGTGAGCTGGGTGTTTGGCGTGGGCTCCAAAGCCCTGTGCT	138.80
Db	1395	AGCTACACCCCTGAGCCAGCAGTGTGAGCTGGGTGTTTGGCGTGGGCTCCAAAGCCCTGTGCT	145.54
QY	1381	TACATGCATGACTGACACCAAGCTGTGTGTCCACCGGGAAGGCCAAAGGACACATGGTGTGC	144.00
Db	1455	TACATGCATGACTGACACCAAGCTGTGTGTCCACCGGGAAGGCCAAAGGACACATGGTGTGC	151.44
QY	1441	CAGACCCGGCACTTCCCTCGGGGCGCATTTGGCACACAGTGTGGGAGAGGCAAGCTCTGCTC	150.00
Db	1515	CAGACCCGGCACTTCCCTCGGGGCGCATTTGGCACACAGTGTGGGAGAGGCAAGCTCTGCTC	157.44
QY	1501	AAAGGGGCGCTGCGTGGAGAGACACAACTCTCAACAAGACACAGGGTGGATGTCTCGGGCC	156.00
Db	1575	AAAGGGGCGCTGCGTGGAGAGACACAACTCTCAACAAGACACAGGGTGGATGTCTCGGGCC	163.44
QY	1561	AAATGGATCCCTATGAGCCCTCTGCGCACATGTGTGGGGGCGTGCACAGCTGGCACGG	162.00
Db	1635	AAATGGATCCCTATGAGCCCTCTGCGCACATGTGTGGGGGCGTGCACAGCTGGCACGG	169.44
QY	1621	AGGCAGTGCACAAACCCCAACCCCTGGCCAAACGGGGGCGTAAGTACTGCGAGAGAGTGAAGGTG	168.00
Db	1695	AGGCAGTGCACAAACCCCAACCCCTGGCCAAACGGGGGCGTAAGTACTGCGAGAGAGTGAAGGTG	175.44
QY	1681	AAATACCGATCTCTGACATCTGAGAGCCCTGCGCCAGTCAAGCTCCGGGAAAGAGCTTCCGG	174.00
Db	1755	AAATACCGATCTCTGACATCTGAGAGCCCTGCGCCAGTCAAGCTCCGGGAAAGAGCTTCCGG	181.44
QY	1741	GAGGAGCAGTGTGAGGCTTTCAACGGCTCAACACACAGCACCAACCGGCTCACTCTGCC	180.00
Db	1815	GAGGAGCAGTGTGAGGCTTTCAACGGCTCAACACACAGCACCAACCGGCTCACTCTGCC	187.44
QY	1801	GTTGCGATAGGGGTGCCCAAGTACTCCGGGCGTGTCTCCCGGGACAAAGTGCATGTACTCTGC	186.00
Db	1875	GTTGCGATAGGGGTGCCCAAGTACTCCGGGCGTGTCTCCCGGGACAAAGTGCATGTACTCTGC	193.44
QY	1861	CGAGCCAAATGGCACTGGACTCTCTATATGTGTGGGACCCCA---GATGGTGGACGGCACG	191.74
Db	1935	CGAGCCAAATGGCACTGGACTCTCTATATGTGTGGGACCCCAAGTGTGGTGTGACGGCACG	199.44
QY	1918	CTGTGCTCTCTGACTGCACCTCCGTTGTGTGCCAAGGCAAGTGCATCAAGGCTGGCTGT	197.74
Db	1995	CTGTGCTCTCTGACTGCACCTCCGTTGTGTGCCAAGGCAAGTGCATCAAGGCTGGCTGT	205.44
QY	1978	GATGGGAACTGGGCTCTCAAGAAAGATTCGACAAAGTGTGGGGTGTGTGGGGAGACAT	203.74
Db	2055	GATGGGAACTGGGCTCTCAAGAAAGATTCGACAAAGTGTGGGGTGTGTGGGGAGACAT	211.44
QY	2038	AAGAGCTGCAAGAAAGTGTGAGACCTTCACCAACCCCATGATGGCTACATATTGCTG	209.74
Db	2115	AAGAGCTGCAAGAAAGTGTGAGACCTTCACCAACCCCATGATGGCTACATATTGCTG	217.44
QY	2098	GTTGCGATCTCCCGCAGGCGGCTCAACATCGACATCCGCGACGCGGTATCAAGGCGCTG	215.74
Db	2175	GTTGCGATCTCCCGCAGGCGGCTCAACATCGACATCCGCGACGCGGTATCAAGGCGCTG	223.44
QY	2158	ATCGGGGATGACACTACTGCTGTGAAGAACAGCCAAAGGCAAGTACCTGTCTCAACGGG	221.74
Db	2235	ATCGGGGATGACACTACTGCTGTGAAGAACAGCCAAAGGCAAGTACCTGTCTCAACGGG	229.44
QY	2218	CATTTCGTGGTGTGCGGGGTGGAGCGGGGACCTGTGTGTGAAGGCAAGTCTGTGGGGTAC	227.74
Db	2295	CATTTCGTGGTGTGCGGGGTGGAGCGGGGACCTGTGTGTGAAGGCAAGTCTGTGGGGTAC	235.44
QY	2278	AGCGGCAAGGGACACAGCGGTGAGAGACCTGACAAAGGCTTCCGCGCCATCCTGAGAGCGGTG	233.74
Db	2355	AGCGGCAAGGGACACAGCGGTGAGAGACCTGACAAAGGCTTCCGCGCCATCCTGAGAGCGGTG	241.44
QY	2338	ACCGTGAAGTCTCTCTCGGGGGAAGATGACACCGCGCCCGGGGTCCGCTACTCTTCTAT	239.74

Db	2415	ACCGGGAGGCTCTCTCGTGGGAGATACACCGCCCGGGCTCGCTACTCTTAT	2474
QY	2398	CTGCCCAAGAGCCTCGGGAGAGACAAGTCTTCATCCC	2436
Db	2475	CTGCCCAAGAGCCTCGGGAGAGACAAGTCTTCATCCC	2513
<p>RESULT 6</p> <p>US-10-163-316-3</p> <p>Sequence 3, Application: .US/10163316</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Kapeller-Libermann, Rosana</p> <p>TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses Therefor</p> <p>FILE REFERENCE: MP101-025PIRNM</p> <p>CURRENT APPLICATION NUMBER: US/10/163,316</p> <p>PRIOR FILING DATE: 2002-06-05</p> <p>PRIOR APPLICATION NUMBER: 60/297,863</p> <p>PRIOR FILING DATE: 2001-06-13</p> <p>NUMBER OF SEQ ID NOS: 10</p> <p>SOFTWARE: FastSeq for Windows Version 4.0</p> <p>SEQ ID NO 3</p> <p>LENGTH: 2469</p> <p>TYPE: DNA</p> <p>ORGANISM: Homo sapiens</p> <p>FEATURE:</p> <p>NAME/KEY: CDS</p> <p>LOCATION: (1)...(2469)</p> <p>US-10-163-316-3</p>			
QY	1	ATGCTTCTGCTGGGCATCTTAACCTTGCTTTCGCGGGGGAACCGCTGAGGCTCTGAG	60
Db	1	ATGCTTCTGCTGGGCATCTTAACCTTGCTTTCGCGGGGGAACCGCTGAGGCTCTGAG	60
QY	61	CCAGAGCGGGAGGTAGTCTCCCATCCGACTGGACCCGGGACATTAAAGGCCGCCGCTAC	120
Db	61	CCAGAGCGGGAGGTAGTCTCCCATCCGACTGGACCCGGGACATTAAAGGCCGCCGCTAC	120
QY	121	TACTGGCGGGTCCCGAGAGCTCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT	180
Db	121	TACTGGCGGGTCCCGAGAGCTCCGGGGATCAGGAGCTCAATTTTCAGATCAGAGCATTT	180
QY	181	CAGAGGACTTTTACCTACACACTGAGCCCGGATGCTCAAGTTTGGCTCCCGCTTCTCC	240
Db	181	CAGAGGACTTTTACCTACACACTGAGCCCGGATGCTCAAGTTTGGCTCCCGCTTCTCC	240
QY	241	ACTGAGCATCTGGGCGTCCCGCTCCAGGGGGCTCACCGGGGGCTCTTACAGACTGCGACGC	300
Db	241	ACTGAGCATCTGGGCGTCCCGCTCCAGGGGGCTCACCGGGGGCTCTTACAGACTGCGACGC	300
QY	301	TGCTTCTATTCTGGGGAGCTGAACGCCGAGCGGAGCTGTTGGCTGCTGAGACCTGTGC	360
Db	301	TGCTTCTATTCTGGGGAGCTGAACGCCGAGCGGAGCTGTTGGCTGCTGAGACCTGTGC	360
QY	361	GGGGGGCTCCGCGGACGCTTTGGCTTACCGAGGGCCCGAGTATGCTAATACCCGCTGCC	420
Db	361	GGGGGGCTCCGCGGACGCTTTGGCTTACCGAGGGCCCGAGTATGCTAATACCCGCTGCC	420
QY	421	AATGCTAGCGGGCGGGGGCGACGGACAAGCCAGGGGGCGACACCTTTCACAGCGCGCG	480
Db	421	AATGCTAGCGGGCGGGGGCGACGGACAAGCCAGGGGGCGACACCTTTCACAGCGCGCG	480
QY	481	GGTGTTCGGGCGGGGCTTCCGAGAGACCCACACTCTTGCTGCGGGGTGGCTCGGGCTGG	540
Db	481	GGTGTTCGGGCGGGGCTTCCGAGAGACCCACACTCTTGCTGCGGGGTGGCTCGGGCTGG	540
QY	541	AACCCGCGCATCTCAAGGGGCTTGAGCCCTTACAGACCGCGGGGGGCTTGGGGAG	600
Db	541	AACCCGCGCATCTCAAGGGGCTTGAGCCCTTACAGACCGCGGGGGGCTTGGGGAG	600

Db 541 AACCCGCCATCTACGGGCCCCCTGACCCCTTACAAGCCGCGGCGGGCTTGGGAG 600
 QY 601 AGTCGTAGCCGGCGAGGCTTGGGGCGCCCAAGCGTTTCGTGTATCCCGGTACTG 660
 Db 601 AGTCGTAGCCGGCGAGGCTTGGGGCGCCCAAGCGTTTCGTGTATCCCGGTACTG 660
 QY 661 GAGAGCGTGGTGGCGAGAGTCAATGTGCAAGTTCCACGGCGGCGACTGGAACAT 720
 Db 661 GAGAGCGTGGTGGCGAGAGTCAATGTGCAAGTTCCACGGCGGCGACTGGAACAT 720
 QY 721 TATGTGTAGCGCTTGGCAAGCGGCGGAGCTCTACGCCCAATCCAGCATCTCAAC 780
 Db 721 TATGTGTAGCGCTTGGCAAGCGGCGGAGCTCTACGCCCAATCCAGCATCTCAAC 780
 QY 781 CCCATCAACATCGTTGTGTCAAGGTGCTCTTTAGAGATCGTGAATCCGGGCCAAG 840
 Db 781 CCCATCAACATCGTTGTGTCAAGGTGCTCTTTAGAGATCGTGAATCCGGGCCAAG 840
 QY 841 GTCACCGGCATATGGGCCCCCTGACGTGCGCAATTTCTGTGCTGGCAGAGAGCTGAAC 900
 Db 841 GTCACCGGCATATGGGCCCCCTGACGTGCGCAATTTCTGTGCTGGCAGAGAGCTGAAC 900
 QY 901 AAAGTAGTGAACAGCACCCGAGTACTGGGAGCACTGCGCATCTTCCACAGGCGAGAC 960
 Db 901 AAAGTAGTGAACAGCACCCGAGTACTGGGAGCACTGCGCATCTTCCACAGGCGAGAC 960
 QY 961 CTGTGTGAGGCGACCACTGTGACACCCCTGGGCAATGGCTGTGTGTGAGTACCATGTGTAC 1020
 Db 961 CTGTGTGAGGCGACCACTGTGACACCCCTGGGCAATGGCTGTGTGTGAGTACCATGTGTAC 1020
 QY 1021 CCCAAGAAACCTGTCTGTGTATGAGAGCATGGGCTTCATGACCTTACCATGCTGCC 1080
 Db 1021 CCCAAGAAACCTGTCTGTGTATGAGAGCATGGGCTTCATGACCTTACCATGCTGCC 1080
 QY 1081 CACGAGCTGGGCGACAGTTCATCATGTGCCCCATGACATGTAAGTCTGTGAGAGAGTG 1140
 Db 1081 CACGAGCTGGGCGACAGTTCATCATGTGCCCCATGACATGTAAGTCTGTGAGAGAGTG 1140
 QY 1141 TTTGGGAAGCTCCGAGCCCAACCATGATGTGCCGACCCCTCATCAGATCGACCTGACC 1200
 Db 1141 TTTGGGAAGCTCCGAGCCCAACCATGATGTGCCGACCCCTCATCAGATCGACCTGACC 1200
 QY 1201 AACCCGTGACGCTGTGAGTGTCCCATATCATCCGACTTCTGTGAGAGCGGCAAGGT 1260
 Db 1201 AACCCGTGACGCTGTGAGTGTCCCATATCATCCGACTTCTGTGAGAGCGGCAAGGT 1260
 QY 1261 GACTGCTCTGTGAGCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
 Db 1261 GACTGCTCTGTGAGCAACCCAGCAAGCCCATCTCCCTGCCCGAGGATCTGCCGGGCGCC 1320
 QY 1321 AGCTACACCTGTGAGCCAGTGTGAGCTGTGGCTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
 Db 1321 AGCTACACCTGTGAGCCAGTGTGAGCTGTGGCTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
 QY 1381 TACATAGAGTGTGACCAAGCTGTGTGCAACCGGGAAGGCGCAAGGAGACAGTGTGTGC 1440
 Db 1381 TACATAGAGTGTGACCAAGCTGTGTGCAACCGGGAAGGCGCAAGGAGACAGTGTGTGC 1440
 QY 1441 CAGACCGGCGACTTCCCTGGGGCGATGAGCAACAGCTGTGGCGAGGCGCAAGCTGTGCTC 1500
 Db 1441 CAGACCGGCGACTTCCCTGGGGCGATGAGCAACAGCTGTGGCGAGGCGCAAGCTGTGCTC 1500
 QY 1501 AAAGGGGCGTGGTGGAGAGACACACCTCAACAGCAGAG 1542
 Db 1501 AAAGGGGCGTGGTGGAGAGACACACCTCAACAGCAGAG 1542

RESULT 7

US-60-297-863-3
 ; Sequence 3, Application US/60297863
 ; GENERAL INFORMATION:
 ; APPLICANT: Kapeller-Libermann, Rosana
 ; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses

; TITLE OF INVENTION: therefor
 ; FILE REFERENCE: MP12001-025P1(M)
 ; CURRENT APPLICATION NUMBER: US/60/297,863
 ; CURRENT FILING DATE: 2001-06-13
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 2469
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(2469)
 US-60-297-863-3
 Query Match 52.3%; Score 1491; DB 73; Length 2469;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGCTTCTGTGGGCGATCCCTTAACCCCTGCTTTCGCGGCGGAACCGCTGAGAGCTGTAG 60
 Db 1 ATGCTTCTGTGGGCGATCCCTTAACCCCTGCTTTCGCGGCGGAACCGCTGAGAGCTGTAG 60
 QY 61 CCAAGCGGAGGAGTACTGCTTCCATCGACTGAGCCGGACATTACGGCCGCTAC 120
 Db 61 CCAAGCGGAGGAGTACTGCTTCCATCGACTGAGCCGGACATTACGGCCGCTAC 120
 QY 121 TACTGGGGGGTCCGAGAGATCCGGGGATGCAAGGATCTTTACATACATCAGACATT 180
 Db 121 TACTGGGGGGTCCGAGAGATCCGGGGATGCAAGGATCTTTACATACATCAGACATT 180
 QY 181 CAGGAGACTTTTACTACCTGACGCGCGAGTGTCTGATTTGGCTCCGCTCTCC 240
 Db 181 CAGGAGACTTTTACTACCTGACGCGCGAGTGTCTGATTTGGCTCCGCTCTCC 240
 QY 241 ACTGAGACTTGGGCTCCCCCTCAAGGGGCTACCGGGGGCTCTTCAAGACTGACGC 300
 Db 241 ACTGAGACTTGGGCTCCCCCTCAAGGGGCTACCGGGGGCTCTTCAAGACTGACGC 300
 QY 301 TGGCTTATTTGGGGAGTGAAGCGCGGAGCGGAGCTGGTGGCTGTGAGCTGTGG 360
 Db 301 TGGCTTATTTGGGGAGTGAAGCGCGGAGCGGAGCTGGTGGCTGTGAGCTGTGG 360
 QY 361 GGGGGGCTCCGGGAGCCCTTGGCTACCGAGGCGCGAGTATGATTAAGCCCGTCC 420
 Db 361 GGGGGGCTCCGGGAGCCCTTGGCTACCGAGGCGCGAGTATGATTAAGCCCGTCC 420
 QY 421 AATGCTAGGGCGCGGGCGGCGAGCGCAACAGCCAGGGCGCACACCTTCTCCAGCGCG 480
 Db 421 AATGCTAGGGCGCGGGCGGCGAGCGCAACAGCCAGGGCGCACACCTTCTCCAGCGCG 480
 QY 481 GGTGTTCCGGGGGGGCTTCCGAGAGACCCACCTCTGCGTGGGGGGTGGGCTGG 540
 Db 481 GGTGTTCCGGGGGGGCTTCCGAGAGACCCACCTCTGCGTGGGGGGTGGGCTGG 540
 QY 541 AACCCCGCATCTCAAGGGCCCTGAGCCCTTACAAGCCCGGCGGGGCTTGGGAG 600
 Db 541 AACCCCGCATCTCAAGGGCCCTGAGCCCTTACAAGCCCGGCGGGGCTTGGGAG 600
 QY 601 AGTCGTAGCGGCGAGGCTGTGGGGCGGCGCAAGCGTTTCTGTATCCCGGTAGCTG 660
 Db 601 AGTCGTAGCGGCGAGGCTGTGGGGCGGCGCAAGCGTTTCTGTATCCCGGTAGCTG 660
 QY 661 GAGAGCTGTGTGGCGGAGAGTCAATGTGCAAGTTCCAGGGGCGGAGCTGGAACAT 720
 Db 661 GAGAGCTGTGTGGCGGAGAGTCAATGTGCAAGTTCCAGGGGCGGAGCTGGAACAT 720
 QY 721 TATGTGTAGCGCTCTGTGCAAGCGGCGGAGCTCTACCGCATCCAGCATCTCAAC 780
 Db 721 TATGTGTAGCGCTCTGTGCAAGCGGCGGAGCTCTACCGCATCCAGCATCTCAAC 780
 QY 781 CCCATCAACATCGTTGTGTCAAGGTGCTCTTTAGAGATCGTGAATCCGGGCCAAG 840
 Db 781 CCCATCAACATCGTTGTGTCAAGGTGCTCTTTAGAGATCGTGAATCCGGGCCAAG 840

```

Db 781 CCATCAACATCGTGTGTGTCAGAGTGCTGCTTCTTAGA..ATGTGATCCGCGCCCAAG 840
QY 841 GTACACGGCAATGGGGCCCTGAGCTGCGCCGCACTTCTGTCTCCGGGAGAAAGACGTGAC 900
Db 841 GTACACGGCAATGGGGCCCTGAGCTGCGCCGCACTTCTGT..CCGTGGAGAAAGACGTGAC 900
QY 901 AAAGTAGTGACAGACACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 960
Db 901 AAAGTAGTGACAGACACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 960
QY 961 CTGTGTGGAGACCAACCACTGTGACACCCCTGGCATGGCT..ATGTGGGTACCATGTGTAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGTGACACCCCTGGCATGGCT..ATGTGGGTACCATGTGTAC 1020
QY 1021 CCCAAGAGAGTGTCTGTCTGTCATTGAGAGATGGGCTTCATCAGCTTCCACCATGTC 1080
Db 1021 CCCAAGAGAGTGTCTGTCTGTCATTGAGAGATGGGCTTCATCAGCTTCCACCATGTC 1080
QY 1081 CACGAGCTGGGCGACGTGTTCACATGCCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1140
Db 1081 CACGAGCTGGGCGACGTGTTCACATGCCCCATGACAAATGTGAAGTCTGTGAGAGGTG 1140
QY 1141 TTGGGAGAGTCCGAGCCAAACCAATGATGTCCCGACCTCATCCAGATCGACCTGCC 1200
Db 1141 TTGGGAGAGTCCGAGCCAAACCAATGATGTCCCGACCTCATCCAGATCGACCTGCC 1200
QY 1201 AACCCCTGTGACCTGTGACGTGTGCTGCATCATCAGCAGCTTCTCTGAGACGCGGACGT 1260
Db 1201 AACCCCTGTGACCTGTGACGTGTGCTGCATCATCAGCAGCTTCTCTGAGACGCGGACGT 1260
QY 1261 GACTGCTCTGTGACCAACCCAGCAAGCCCATCTCCCTGCCGAGATCTGCCGCGGCC 1320
Db 1261 GACTGCTCTGTGACCAACCCAGCAAGCCCATCTCCCTGCCGAGATCTGCCGCGGCC 1320
QY 1321 AGCTACACCTGAGCAGCAGTGTGAGTGGGCTTTGGGGGGCTCCCAAGCCCTGTCT 1380
Db 1321 AGCTACACCTGAGCAGCAGTGTGAGTGGGCTTTGGGGGGCTCCCAAGCCCTGTCT 1380
QY 1381 TACATCAGTACTGACCAAGCTGTGTGTGACCGGAGAGGCCAAGGAGACATGTGTGTC 1440
Db 1381 TACATCAGTACTGACCAAGCTGTGTGTGACCGGAGAGGCCAAGGAGACATGTGTGTC 1440
QY 1441 CACACCCGCGCACTTCCCTGGGGCCGATGTGACAGCTGTGTGGAGGCAAGCTCTGCTC 1500
Db 1441 CACACCCGCGCACTTCCCTGGGGCCGATGTGACAGCTGTGTGGAGGCAAGCTCTGCTC 1500
QY 1501 AAAGGGGCTGTGCTGTGAGAGACACAACTCAACAAAGCACAGG 1542
Db 1501 AAAGGGGCTGTGCTGTGAGAGACACAACTCAACAAAGCACAGG 1542

RESULT 8
US-10-163-316-1
; Sequence 1, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MP101-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:

```

```

; NAME/KEY: misc.feature
; LOCATION: (1)-(2940)
; OTHER INFORMATION: n = A,T,C or G
US-10-163-316-1

Query Match
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGTGCGGACCTTAACCCCTGCTTTCGCGGCGCAACCGCTGAGGCTGTGAG 60
Db 472 ATGCTTCTGTGCGGACCTTAACCCCTGCTTTCGCGGCGCAACCGCTGAGGCTGTGAG 531
QY 61 CCAGAGCGGAGAGTGTGTTCCCATCCGATGAGACCCGAGCATTTACGCGCCGCTAC 120
Db 532 CCAGAGCGGAGAGTGTGTTCCCATCCGATGAGACCCGAGCATTTACGCGCCGCTAC 591
QY 121 TACTGGCGGGGTCCGAGAGACTCCGCGGATGAGGAGTCAATTTTCAGATCAGACATTT 180
Db 592 TACTGGCGGGGTCCGAGAGACTCCGCGGATGAGGAGTCAATTTTCAGATCAGACATTT 651
QY 181 CAGAGAGACTTTTACCTACACCTGACGCGGAGATGCTAGTTCTTGGCTCCGCTCTCC 240
Db 652 CAGAGAGACTTTTACCTACACCTGACGCGGAGATGCTAGTTCTTGGCTCCGCTCTCC 711
QY 241 ACTGAGCATTTGGGGCTCCCTCCAGGGGCTCACCGGGGGCTCTTACAGACTTCGACGC 300
Db 712 ACTGAGCATTTGGGGCTCCCTCCAGGGGCTCACCGGGGGCTCTTACAGACTTCGACGC 771
QY 301 TGCCTTCTATTTGGGGAGTGAAGCCGAGCGGACTGTTCGCTGCTGTGAGCTGTGC 360
Db 772 TGCCTTCTATTTGGGGAGTGAAGCCGAGCGGACTGTTCGCTGCTGTGAGCTGTGC 831
QY 361 GGGGGGCTTCGCGGAGCTTTGGCTACCGAGGCGCGGAGTGTGTCATTTAGCCGCTGCC 420
Db 832 GGGGGGCTTCGCGGAGCTTTGGCTACCGAGGCGCGGAGTGTGTCATTTAGCCGCTGCC 891
QY 421 AATGTACGCGCGCGCGCGCAGCGCAAGCAGGAGGCGGAGCAACCTTCTACAGCGCGG 480
Db 892 AATGTACGCGCGCGCGCGCAGCGCAAGCAGGAGGCGGAGCAACCTTCTACAGCGCGG 951
QY 481 GGTGTTCCGGGCGGGCTTTCGCGAGACCCACCTCTGCTGCGGGGTGGCTCGGGCTGG 540
Db 952 GGTGTTCCGGGCGGGCTTTCGCGAGACCCACCTCTGCTGCGGGGTGGCTCGGGCTGG 1011
QY 541 AACCCCGCATCTACAGGCGCTTGACCTTACAAAGCCGCGGCGGGCTTGGGGAG 600
Db 1012 AACCCCGCATCTACAGGCGCTTGACCTTACAAAGCCGCGGCGGGCTTGGGGAG 1071
QY 601 AGCTGTAGCCGGCGCAGGCTGTGGCGCGCCAAAGGTTTGTGTATTCGCCGAGTACGTG 660
Db 1072 AGCTGTAGCCGGCGCAGGCTGTGGCGCGCCAAAGGTTTGTGTATTCGCCGAGTACGTG 1131
QY 661 GAGAGCGTGTGTGTGCGCGAGTCAATGTGTCAAGTTCCAGCGCGGAGCTTGGAAAT 720
Db 1132 GAGAGCGTGTGTGTGCGCGAGTCAATGTGTCAAGTTCCAGCGCGGAGCTTGGAAAT 1191
QY 721 TATCTGTGAGCGTGTGCGCAAGCGCGGCGGAGATCTACCGCCATCCAGCATCTTAAC 780
Db 1192 TATCTGTGAGCGTGTGCGCAAGCGCGGCGGAGATCTACCGCCATCCAGCATCTTAAC 1251
QY 781 CCCATCAACATCTTGTGTGTCAGAGTCTGCTTGTAGAGATGTGATCCGCGGCGCAAG 840
Db 1252 CCCATCAACATCTTGTGTGTCAGAGTCTGCTTGTAGAGATGTGATCCGCGGCGCAAG 1311
QY 841 GTCAACGGCAATGCGGCGCTGTGACGCTGCGCAACTTGTGTGCTGTGAGAGAGTGAAC 900
Db 1312 GTCAACGGCAATGCGGCGCTGTGACGCTGCGCAACTTGTGTGCTGTGAGAGAGTGAAC 1371
QY 901 AAAGTAGTGACAGACACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 960
Db 1372 AAAGTAGTGACAGACACCCCGAGTACTGGGACACTGCC..ATCCTCTTTCACACAGCAGAC 1431

```



```
Db 1612 TTTGGGAGGCTCCGAGGCAACACATGATGTCCCCACCCCTCATCCAGATCGACCTGGCC 1671
QY 1201 AACCCCTGTCAGCTGTCAGTGTGTCATCATCAGCAGCTCTCTGTCAGCGGGGCGCT 1260
Db 1672 AACCCCTGGTCAGCTGTCAGTGTGTCATCATCAGCAGCTCTCTGTCAGCGGGGCGCT 1731
QY 1261 GACTGCTCTGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1320
Db 1732 GACTGCTCTGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1791
QY 1321 AGCTACACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1380
Db 1792 AGCTACACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1851
QY 1381 TACATGACGACTGACACCAAGCTGTGTGACCGGAGGAGGAGCAGATGCTGTGC 1440
Db 1852 TACATGACGACTGACACCAAGCTGTGTGACCGGAGGAGGAGCAGATGCTGTGC 1911
QY 1441 CAGACCCGCGCACTTCCCTGGGCGCATGGACACGCTGTGGGAGGAGCAGCTGTGC 1500
Db 1912 CAGACCCGCGCACTTCCCTGGGCGCATGGACACGCTGTGGGAGGAGCAGCTGTGC 1971
QY 1501 AAGGGGCGCTGCTGGAGAGACCAACCTCAGACAGC 1542
Db 1972 AAGGGGCGCTGCTGGAGAGACCAACCTCAGACAGC 2013

RESULT 10
US-10-093-463-27
; Sequence 27, Application US/10093463
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zernusen, Bryan
; APPLICANT: Tcherny, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
; FILE REFERENCE: 21402-290A (Cura 590A7)
; CURRENT APPLICATION NUMBER: US/10/093,463
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
```

```
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 2804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(2797)
US-10-093-463-27

Query Match 52.2%; Score 1488; DB 39; Length 2804;
Best Local Similarly 99.9%; Pred. No. 0;
Matches 1538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCTTAACCTGCTTTCGCGGGGAGACCGCTGAGCTCTGAG 60
Db 25 ATGCTTCTGCTGGGATCTTAACCTGCTTTCGCGGGGAGACCGCTGAGCTCTGAG 84
QY 61 CCAGAGCGGAGTAGTGTGTTCCCATCCGACTGACCCCGGACATTACGGCGCGCTAG 120
Db 85 CCAGAGCGGAGTAGTGTGTTCCCATCCGACTGACCCCGGACATTACGGCGCGCTAG 144
QY 121 TACTGGCGGGGTCGCCGAGGACTCGGGGATCAGGACTATTTTCAGATCAGCATTT 180
Db 145 TACTGGCGGGGTCGCCGAGGACTCGGGGATCAGGACTATTTTCAGATCAGCATTT 204
QY 181 CAGAGGACCTTTACCTACACCTGAGCGCGGATGCTAGTCTTGGTCCCGCTTCTCC 240
Db 205 CAGAGGACCTTTACCTACACCTGAGCGCGGATGCTAGTCTTGGTCCCGCTTCTCC 264
QY 241 ACTGAGCATCTGGGCTGCCCTTCACAGGGGCTCACCGGGGCTCTTCAGACTGCGACGC 300
Db 265 ACTGAGCATCTGGGCTGCCCTTCACAGGGGCTCACCGGGGCTCTTCAGACTGCGACGC 324
QY 301 TGTCTTATTTCTGGGAGCGTGAACCGGACCGGACTCTTGGTGTGTGAGCTGTGC 360
Db 325 TGTCTTATTTCTGGGAGCGTGAACCGGACCGGACTCTTGGTGTGTGAGCTGTGC 384
QY 361 GGGGGGCTCCGCGGAGCTTGGCTACCGAGGGCGGAGATGATTAACCCGCTGGCC 420
Db 385 GGGGGGCTCCGCGGAGCTTGGCTACCGAGGGCGGAGATGATTAACCCGCTGGCC 444
```


Query Match	44.28;	Score 1261;	DB 39;	Length 2297;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 131;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0

```

QY 2322 CATCTGGAGCGCGTGCACCGGTGAGAGTCTCTCTCCGTGGGGAACATACACCGCCCTGGGT 2381
Db 1761 CATCTGGAGCGCGTGCACCGGTGAGAGTCTCTCTCCGTGGGGAACATACACCGCCCTGGGT 1820
QY 2382 CCGTACTCTCTTATCTATCTGCCAAAGAGCCTCGGAGAGACAGTCTCTATCTCAAGAGA 2441
Db 1821 CCGTACTCTCTTATCTATCTGCCAAAGAGCCTCGGAGAGACAGTCTCTATCTCAAGAGA 1880
QY 2442 CCCCCGGGGACCTCTGTCTTGACAAACAGCGTCTAGCTCTCCAAACAGGTGAGCA 2501
Db 1881 CCCCCGGGGACCTCTGTCTTGACAAACAGCGTCTAGCTCTCCAAACAGGTGAGCA 1940
QY 2502 GCCGAGACAGAGGCCCTCTGCACGCTGGGTGGCTGGGACGTGGGGGCCCGTGCCTCGCGAG 2561
Db 1941 GCCGAGACAGAGGCCCTCTGCACGCTGGGTGGCTGGGACGTGGGGGCCCGTGCCTCGCGAG 2000
QY 2562 CTGCGGCGAGTGGCCTGCAGAAAGCGGGCGGTGAGCTGTGGGGGCTCCGCCGGGACAGCGCAC 2621
Db 2001 CTGCGGCGAGTGGCCTGCAGAAAGCGGGCGGTGAGCTGTGGGGGCTCCGCCGGGACAGCGCAC 2060
QY 2622 GGTCTCTCTCTGTATGATGAGCCCATGGGCCGGTGGAGACACAAGCCTGGGGGGACCTCTG 2681
Db 2061 GGTCTCTCTCTGTATGATGAGCCCATGGGCCGGTGGAGACACAAGCCTGGGGGGACCTCTG 2120
QY 2682 CCCCACCTGGGAGCTCAGCGCGCTGTGTACCCCTGCTCCAAAGAGCTGGCGCGGGGAGATTCA 2741
Db 2121 CCCCACCTGGGAGCTCAGCGCGCTGTGTACCCCTGCTCCAAAGAGCTGGCGCGGGGAGATTCA 2180
QY 2742 GAGGGGCTCATCAAGTGTGTGGGCCGAGAGGCCGGGCTGTGTGGCCCGGGAGCACTGCANA 2801
Db 2181 GAGGGGCTCATCAAGTGTGTGGGCCGAGAGGCCGGGCTGTGTGGCCCGGGAGCACTGTGCANA 2240
QY 2802 CTGGACCGCAAGCCCGCAGGAGCTGAGCTTCTCGTCTCGAGAGGCCGTGTGTGA 2853
Db 2241 CTGGACCGCAAGCCCGCAGGAGCTGAGCTTCTCGTCTCGAGAGGCCGTGTGTGA 2292

RESULT 12
US-60-212-656-762
; Sequence 762, Application US/60212656
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000673
; CURRENT APPLICATION NUMBER: US/60/212,656
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 762
; LENGTH: 3400
; TYPE: DNA
; ORGANISM: HUMAN
US-60-212-656-762

Query Match 44.2%; Score 1261; DB 65; Length 3400;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1542 GGTGATGATGTTCTGGGCGCAAAATGGATCCCATATGGCCCTGCTGGCCACATGTGTTGG 1601
Db 2089 GGTGATGATGTTCTGGGCGCAAAATGGATCCCATATGGCCCTGCTGGCCACATGTGTTGG 2148
QY 1602 GGGGCTGACACTGGCCAGAGGAGCATGTGCACCAACCCACCCCTGCGCAAGSGGGSCAAGTA 1661
Db 2149 GGGGCTGACACTGGCCAGAGGAGCATGTGCACCAACCCACCCCTGCGCAAGSGGGSCAAGTA 2208
QY 1662 CTGCGAGGAGTGAAGGTAAATACGATCTCGCAATCTGGAGGACCTGCCCCACACTCAGC 1721
Db 2209 CTGCGAGGAGTGAAGGTAAATACGATCTCGCAATCTGGAGGACCTGCCCCACACTCAGC 2268
QY 1722 CTCGCGAAGAGCTTCCGAGAGAGCATGTGTAGGCTTTCAACAGGCTACACACACACAC 1781

```

QY	1542	GGTGGATGTTCTTGCGCCAAATGAGATCCCTATGGCCCTCGTGGCGCATGTGTGG	1601
QY	1542	GGTGGATGTTCTTGCGCCAAATGAGATCCCTATGGCCCTCGTGGCGCATGTGTGG	1601
Db	2448	GGTGGATGTTCTTGCGCCAAATGAGATCCCTATGGCCCTCGTGGCGCATGTGTGG	2507
QY	1602	GGGCGTGCAGCTGGCGCAGAGAGCAGTGCACCAACCCACCCTGCGCAAGGGGGCAGTA	1661
Db	2508	GGGCGTGCAGCTGGCGCAGAGAGCAGTGCACCAACCCACCCTGCGCAAGGGGGCAGTA	2567
QY	1662	CTGCGAGGAGTGAAGGTTGAATACGATCCTTGCAATCTGAGAGCCTGCCCCACGCTCAG	1721
Db	2568	CTGCGAGGAGTGAAGGTTGAATACGATCCTTGCAATCTGAGAGCCTGCCCCACGCTCAG	2627
QY	1722	CTCCGGAAGAAGCTCCGGGGAGAGCAGTGTGAGAGCTTCAACGCGCTAACACACAGCAGC	1781
Db	2628	CTCCGGAAGAAGCTCCGGGGAGAGCAGTGTGAGAGCTTCAACGCGCTAACACACAGCAGC	2687
QY	1782	CAACCGGCTCACTCCGCCGTGGCAGTGGGATGGCCCAAGTACTCCGGCGTGTCTCCCGGGA	1841
Db	2688	CAACCGGCTCACTCCGCCGTGGCAGTGGGATGGCCCAAGTACTCCGGCGTGTCTCCCGGGA	2747
QY	1842	CAAGTGAAGTCACTCTGCGCAGGAGGCAAGGACATGGCTACTTGTATGTCTGGACCCAA	1901
Db	2748	CAAGTGAAGTCACTCTGCGCAGGAGGCAAGGACATGGCTACTTGTATGTCTGGACCCAA	2807
QY	1902	GGTGGTGAAGCGCAAGCTGTGCTCTCTGACTCCACCTCCGCTGTGTGTCCAAAGCAGTG	1961
Db	2808	GGTGGTGAAGCGCAAGCTGTGCTCTCTGACTCCACCTCCGCTGTGTGTGTCCAAAGCAGTG	2867
QY	1962	CATCAAGCTGGCTGTGTATGAGAACTGGGCTCCAAAGAAAGTGTGACAAAGTGTGGGT	2021
Db	2868	CATCAAGCTGGCTGTGTATGAGAACTGGGCTCCAAAGAAAGTGTGACAAAGTGTGGGT	2927
QY	2022	GGTGGGGGAGACATTAAGAGCTCAAGAAAGTGAAGTGAAGTCTTTCACCAAGCCATGA	2081
Db	2928	GGTGGGGGAGACATTAAGAGCTCAAGAAAGTGAAGTGAAGTCTTTCACCAAGCCATGA	2987
QY	2082	TGGCTACATTTTGTGTGGCCATCCCGCAGGCGCTCAAGATGACATCCGCCAGG	2141
Db	2988	TGGCTACATTTTGTGTGGCCATCCCGCAGGCGCTCAAGATGACATCCGCCAGG	3047
QY	2142	CGGTTACAAAGGCGTGAATGGGGATGACACTACTCTGCTCAAGAGAGCCAAAGCAA	2201

```

Query Match:      38.2%; Score 1091; DB 36; Length 1104;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTTCGTGGCGACTTAACCGCGGCTTCCCCGGGCGAACCCGCTGAGGCTCTGAG 60
|||||

```

Db	1	ATGCTTCTGCTGGGATCTCTAAACCGTGGCTTTGGCCGGGGGAACCGGTGGAGGCTCTGAG	60
QY	61	CCAGAGCCGGAGGATGCTGCTTCCATCCGACTGAGACCGGACATTAACGGCCGGCTAC	120
Db	61	CCAGAGCCGGAGGATGCTGCTTCCATCCGACTGAGACCGGACATTAACGGCCGGCTAC	120
QY	121	TACTGGCGGGGTCCCGAGAGCTCCGGGGATAGGAGACTCAATTTTCAGATCAAGCATTT	180
Db	121	TACTGGCGGGGTCCCGAGAGCTCCGGGGATAGGAGACTCAATTTTCAGATCAAGCATTT	180
QY	181	CAGGAGGACTTTTACCTACACCTGACGGCGGATGCTCAGTTCTTGCTCCGCTTCTCC	240
Db	181	CAGGAGGACTTTTACCTACACCTGACGGCGGATGCTCAGTTCTTGCTCCGCTTCTCC	240
QY	241	ACTGAGCATCTGGGGGTCCCGCTCCCAAGGGGTCAACGGGGGGCTTTCAGACTGCGAGCG	300
Db	241	ACTGAGCATCTGGGGGTCCCGCTCCCAAGGGGTCAACGGGGGGCTTTCAGACTGCGAGCG	300
QY	301	TGCTTCTATTCTGGGGAGCTGAAGCGCGAGCGGACACTGTTCCCTGCTGAGACTCTGTGC	360
Db	301	TGCTTCTATTCTGGGGAGCTGAAGCGCGAGCGGACACTGTTCCCTGCTGAGACTCTGTGC	360
QY	361	GGGGGGGCTCCCGGAGGCTTTGGCTACCGAGGGCGCGAGTATGTCATTAAGCCGTCGCC	420
Db	361	GGGGGGGCTCCCGGAGGCTTTGGCTACCGAGGGCGCGAGTATGTCATTAAGCCGTCGCC	420
QY	421	AATGCTACGCGCGCGCGGCGACGCAACGACGAGGGCGCACACCTTCTCCAGCCGCGG	480
Db	421	AATGCTACGCGCGCGCGGCGACGCAACGACGAGGGCGCACACCTTCTCCAGCCGCGG	480
QY	481	GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGCGGGGTGGCCTTGGGCTGG	540
Db	481	GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTCGCTGCGGGGTGGCCTTGGGCTGG	540
QY	541	AACCCCGCATCTCAACGGGGCCCTGGACCTTACAAGCCGGGGGGCGGGCTTCCGGGAG	600
Db	541	AACCCCGCATCTCAACGGGGCCCTGGACCTTACAAGCCGGGGGGCGGGCTTCCGGGAG	600
QY	541	AACCCCGCATCTCAACGGGGCCCTGGACCTTACAAGCCGGGGGGCGGGCTTCCGGGAG	600
Db	541	AACCCCGCATCTCAACGGGGCCCTGGACCTTACAAGCCGGGGGGCGGGCTTCCGGGAG	600
QY	601	AGTGTACCGGCGCAGGTCTGGGCGGCGCCAAAGGTTTCTGTCTATCCCGGGTACGTG	660
Db	601	AGTGTACCGGCGCAGGTCTGGGCGGCGCCAAAGGTTTCTGTCTATCCCGGGTACGTG	660
QY	661	GAGAGCGGTGTGTGTGGGAGAGATCAATGCTCAAGTTCCAGCGCGCGGACCTGGACAT	720
Db	661	GAGAGCGGTGTGTGTGGGAGAGATCAATGCTCAAGTTCCAGCGCGCGGACCTGGACAT	720
QY	721	TATCTGTGAGCGCTGTGGCAACGGCGCGGACTGTACCGGCATCCGACATTCCTCAAC	780
Db	721	TATCTGTGAGCGCTGTGGCAACGGCGCGGACTGTACCGGCATCCGACATTCCTCAAC	780
QY	781	CCCATCAACATCTGTTGTGTGTCAGAGTGTCTCTTTAGAGATGTGATCTCCGGGCCCAAG	840
Db	781	CCCATCAACATCTGTTGTGTGTCAGAGTGTCTCTTTAGAGATGTGATCTCCGGGCCCAAG	840
QY	841	GTCACCGGCATTTGGGGCCCTGAGCGTGGCGCAATTCGTGGGCGCGAGAGAAGAGCTGAC	900
Db	841	GTCACCGGCATTTGGGGCCCTGAGCGTGGCGCAATTCGTGGGCGCGAGAGAAGAGCTGAC	900
QY	901	AAAGTAGTGACAGCAACCCCGAGTACTGGGACACTGCCATCCTCTTCAACGAGCGAGAC	960
Db	901	AAAGTAGTGACAGCAACCCCGAGTACTGGGACACTGCCATCCTCTTCAACGAGCGAGAC	960
QY	961	CTGTGTGGAGCCACACTCTGTGACACCTCTGGGATGGGTGATGTGGGTACATGTGTGAC	1020
Db	961	CTGTGTGGAGCCACACTCTGTGACACCTCTGGGATGGGTGATGTGGGTACATGTGTGAC	1020
QY	1021	CCCAAGAGAAAGCTCTCTGTCAATTTGAGAGAGATGGGCTTCAATCAAGCTTCAACCTGCC	1080
Db	1021	CCCAAGAGAAAGCTCTCTGTCAATTTGAGAGAGATGGGCTTCAATCAAGCTTCAACCTGCC	1080
QY	1081	CACGAGCTGGG 1091	
Db	1081	CACGAGCTGGG 1091	

```
RESULT 15
US-09-965-631-1
; Sequence 1, Application US/09965631
; GENERAL INFORMATION:
; APPLICANT: Fiddale, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 966
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-965-631-1
```

```
Query Match      33.6%; Score 959; DB 36; Length 966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCTTGTGCTGGGATCCTAACCCCTGGCTTTGCGCGGGAGACCGCTGAGGCTCTGAG 60
   |||||||
DB 1 ATGCTTGTGCTGGGATCCTAACCCCTGGCTTTGCGCGGGAGACCGCTGAGGCTCTGAG 60

QY 61 CCAGAGCGGAGGATGCTGCTTCCATCCAGACTGAGACCGGACATTAAGCGCGCGCTAC 120
   |||||||
DB 61 CCAGAGCGGAGGATGCTGCTTCCATCCAGACTGAGACCGGACATTAAGCGCGCGCTAC 120

QY 121 TACTGGCGGGGTCCCGAGGACTCCGGGGATCAGGACTATTTTTCAGATCAGACATT 180
   |||||||
DB 121 TACTGGCGGGGTCCCGAGGACTCCGGGGATCAGGACTATTTTTCAGATCAGACATT 180

QY 181 CAGGAGACTTTTACCTACACCTGAGCGCGGATGCTCAGTTTGGCTCCCGCTCTCTCC 240
   |||||||
DB 181 CAGGAGACTTTTACCTACACCTGAGCGCGGATGCTCAGTTTGGCTCCCGCTCTCTCC 240

QY 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGGCTTTTCAGACTGCGAGCG 300
   |||||||
DB 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGGCTTTTCAGACTGCGAGCG 300

QY 301 TGCCTTATTTCTGGGGAGCGAAGCGCGGACCTGCTGCTGTGAGCTGTGC 360
   |||||||
DB 301 TGCCTTATTTCTGGGGAGCGAAGCGCGGACCTGCTGCTGTGAGCTGTGC 360

QY 361 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGCGCGAGTAGTATTCATTAAGCCGCTGCC 420
   |||||||
DB 361 GGGGGGCTCCGCGAGGCTTTGGCTACCGAGCGCGAGTAGTATTCATTAAGCCGCTGCC 420

QY 421 AATGCTAGCGCGCGCGGCGGAGCGGACGACAGCGAGGCGCACACCTTCTCCAGCGCGG 480
   |||||||
DB 421 AATGCTAGCGCGCGCGGCGGAGCGGACGACAGCGAGGCGCACACCTTCTCCAGCGCGG 480

QY 481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCTCGGGCTGG 540
   |||||||
DB 481 GGTGTTCCGGGGGGGCTTCCGGAGACCCACCTCTGCTGCGGGGTGGCTCGGGCTGG 540

QY 541 AACCCGCGCATCTACGCGGCTTGAGACCTTACAGCGCGGCGGGGCGGCTTCGGGGAG 600
   |||||||
DB 541 AACCCGCGCATCTACGCGGCTTGAGACCTTACAGCGCGGCGGGGCGGCTTCGGGGAG 600

QY 601 AGTGTAGCGCGCGGAGGCTTGGGCGCGGCAAGCGCTTGTGCTATCCCGGCTACGTG 660
   |||||||
DB 601 AGTGTAGCGCGCGGAGGCTTGGGCGCGGCAAGCGCTTGTGCTATCCCGGCTACGTG 660

QY 661 GAGACGCTGGTGTGCGGCGGAGGATCAATGATTCAGGCTGAGGCGGAGCTGGAACAT 720
   |||||||
DB 661 GAGACGCTGGTGTGCGGCGGAGGATCAATGATTCAGGCTGAGGCGGAGCTGGAACAT 720
```

```
QY 721 TATCGCTGACGCTGTGGCAACGGCGGAGCTTACCGCCATCCAGATCTCTCAAC 780
   |||||||
DB 721 TATCGCTGACGCTGTGGCAACGGCGGAGCTTACCGCCATCCAGATCTCTCAAC 780

QY 781 CCATCAACATCGTTGTGTCAGAGGTGCTGCTTTAGAGATCGTAGCTCCGGGCCAG 840
   |||||||
DB 781 CCATCAACATCGTTGTGTCAGAGGTGCTGCTTTAGAGATCGTAGCTCCGGGCCAG 840

QY 841 GTCACCGGCAATGGGCGCTTACGCTGCGGCACTTCTGTGCTGGGAGAGAGCTGAAC 900
   |||||||
DB 841 GTCACCGGCAATGGGCGCTTACGCTGCGGCACTTCTGTGCTGGGAGAGAGCTGAAC 900

QY 901 AAAGTAGTAGCAAGACACCCGAGTACGAGACACTGCATCTCTTCACAGGACAGA 959
   |||||||
DB 901 AAAGTAGTAGCAAGACACCCGAGTACGAGACACTGCATCTCTTCACAGGACAGA 959
```

```
Search completed: May 16, 2003, 05:38:15
Job time : 5618 secs
```

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 03:53:22 ; Search time 12839 Seconds
(without alignments)
1149.891 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853
Sequence: 1 atgctctgctgctgcatcctc.....gctcctgagcgcgtgctga 2853

Scoring table: OLIGO NWC
Gapop 60.0 , Gapext 60.0

Searched: 6404235 seqs, 2587356060 residues

Word size : 24

Total number of hits satisfying chosen parameters: 24

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2802	98.2	2853	8 US-10-391-364-76	Sequence 76, Appl
2	2802	98.2	2853	8 US-10-391-364-78	Sequence 78, Appl
3	2802	98.2	2853	8 US-10-170-235-10538	Sequence 10538, A
4	2802	98.2	2853	10 US-60-466-411-7	Sequence 7, Appl
5	2469	86.5	2853	11 US-60-453-135-7278	Sequence 7278, Ap
6	2469	86.5	2853	11 US-60-453-050-7278	Sequence 32, Appl
7	2196	77.0	2930	9 US-10-311-035-32	Sequence 77890, A
8	150	5.3	201	11 US-60-453-135-77889	Sequence 77890, A
9	150	5.3	201	11 US-60-453-135-77891	Sequence 77891, A
10	150	5.3	201	11 US-60-453-135-77892	Sequence 77892, A
11	150	5.3	201	11 US-60-453-135-77893	Sequence 77893, A
12	150	5.3	201	11 US-60-453-135-77894	Sequence 77894, A
13	150	5.3	201	11 US-60-453-050-77889	Sequence 77889, A
14	150	5.3	201	11 US-60-453-050-77890	Sequence 77890, A
15	150	5.3	201	11 US-60-453-050-77891	Sequence 77891, A
16	150	5.3	201	11 US-60-453-050-77892	Sequence 77892, A
17	150	5.3	201	11 US-60-453-050-77893	Sequence 77893, A
18	127	4.5	178	11 US-60-453-135-77892	Sequence 77892, A
19	127	4.5	178	11 US-60-453-050-77892	Sequence 77892, A
20	126	4.4	177	11 US-60-453-135-77895	Sequence 77895, A
21	126	4.4	177	11 US-60-453-050-77895	Sequence 77895, A
22	100	3.5	115	11 US-60-453-135-77896	Sequence 77896, A

23 100 3.5 115 11 US-60-453-050-77896 Sequence 77896, A
24 74 2.6 5714 9 US-10-144-771-12354 Sequence 12354, A

ALIGNMENTS

```
RESULT 1
US-10-391-364-76
; Sequence 76, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Welch, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352.
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MP103-0150M1N
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2853)
; US-10-391-364-76
Query Match 98.2% Score 2802; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 ATGCTTGTGCTGGGAGTCCTAACCCTGGCTTCCCGGGGAGAACCGCTGAGGCTTGAG 60
DB 1 ATGCTTGTGCTGGGAGTCCTAACCCTGGCTTCCCGGGGAGAACCGCTGAGGCTTGAG 60
QY 61 CCAAGAGCGGAGGTAGTGTCTTCCATCCGACTGAGACCCGGAGATTAAAGGCGCCGCTAC 120
DB 61 CCAAGAGCGGAGGTAGTGTCTTCCATCCGACTGAGACCCGGAGATTAAAGGCGCCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGAGTCAGGACTCAATTTTTCAGATCAGACATTT 180
DB 121 TACTGGCGGGTCCCGAGGACTCCGGGAGTCAGGACTCAATTTTTCAGATCAGACATTT 180
QY 181 CAGAGGAGCTTTTACCTACACCTGACGCCGAGATGCTCAGTTCTTGCTCCCGCTTCTCC 240

|||||
Db 181 CAGAGGACTTTACCTACACCTGACCGCGATCTCAGTTCTGGCTCCCGCTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTAGACCTGCGACGC 300
Db 241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGCTCTTAGACCTGCGACGC 300
QY 301 TGTCTATTCTGGGAGAGTGAACGCGGAGCTGTTGGCTGTGTGAGCTGTGCTG 360
Db 301 TGTCTATTCTGGGAGAGTGAACGCGGAGCTGTTGGCTGTGTGAGCTGTGCTG 360
QY 361 GGGGGGCTCCGGGAGCTTTGGCTACCGAGGGCGCGAGTATGTCAATTAGCCGCTGCC 420
Db 361 GGGGGGCTCCGGGAGCTTTGGCTACCGAGGGCGCGAGTATGTCAATTAGCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGGCGGAGCGACGACAGCGGAGCGACACCTTCCTCAAGCGCGG 480
Db 421 AATGCTAGCGCGCGGGGCGGAGCGACGACAGCGGAGCGACACCTTCCTCAAGCGCGG 480
QY 481 GGTGTCCGGGCGGGGCTTCGCGAGAGCCGACCTCTGCTGCGGGGTGGGCTGGGCTGG 540
Db 481 GGTGTCCGGGCGGGGCTTCGCGAGAGCCGACCTCTGCTGCGGGGTGGGCTGGGCTGG 540
QY 541 AATCCCGCATCTCAGGGGCGCTTGAGACCGCGTGGGGGCGGGGCTTCGCGGAG 600
Db 541 AATCCCGCATCTCAGGGGCGCTTGAGACCGCGTGGGGGCGGGGCTTCGCGGAG 600
QY 601 AGTGTAGCGCGCGGAGGCTGCGGCGCGCAAGGTTTCTGTATCTCCGCGGTACGCTG 660
Db 601 AGTGTAGCGCGCGGAGGCTGCGGCGCGCAAGGTTTCTGTATCTCCGCGGTACGCTG 660
QY 661 GAGAGCTGTGTGTCGCGGAGAGTCAATGTGTCAGTTCCAGGCGCGGAGCTTGGAACT 720
Db 661 GAGAGCTGTGTGTCGCGGAGAGTCAATGTGTCAGTTCCAGGCGCGGAGCTTGGAACT 720
QY 721 TATGTGTCAGGCTGTGGGCAAGGGCGGAGCTACCGGCACTCCACATCTCTCAAC 780
Db 721 TATGTGTCAGGCTGTGGGCAAGGGCGGAGCTACCGGCACTCCACATCTCTCAAC 780
QY 781 CCCATCAACATCTGTGTGTCAGAGTGTCTGTCTTAGAGATCGTGACTCCGGGCCAAG 840
Db 781 CCCATCAACATCTGTGTGTCAGAGTGTCTGTCTTAGAGATCGTGACTCCGGGCCAAG 840
QY 841 GTACACGGCAATGTGGGCGCTGAGCGTGCACAATCTTGTGTGTGCGGAGGAGCTGAAC 900
Db 841 GTACACGGCAATGTGGGCGCTGAGCGTGCACAATCTTGTGTGTGCGGAGGAGCTGAAC 900
QY 901 AAGAGTGTACAGCAACCCGAGTACTGGGAGCACTGCAATCTCTTACACAGGAGAGAC 960
Db 901 AAGAGTGTACAGCAACCCGAGTACTGGGAGCACTGCAATCTCTTACACAGGAGAGAC 960
QY 961 CTGTGTGAGGCGACACCTGTGACACCTGTGGCATGTGCTATGTGGTACTATGTGTGAC 1020
Db 961 CTGTGTGAGGCGACACCTGTGACACCTGTGGCATGTGCTATGTGGTACTATGTGTGAC 1020
QY 1021 CCCAAGAGAACTGCTGTGTATTTAGAGAGATGGGCTTCATCAGCTTCACACCTGCC 1080
Db 1021 CCCAAGAGAACTGCTGTGTATTTAGAGAGATGGGCTTCATCAGCTTCACACCTGCC 1080
QY 1081 CACAGGCTGGGCGACGTTTCAACATGCCCATGACATGGAAGTCTGTGAGGAGGTG 1140
Db 1081 CACAGGCTGGGCGACGTTTCAACATGCCCATGACATGGAAGTCTGTGAGGAGGTG 1140
QY 1141 TTTGGGAGGCTCCGAGCGCAACACATGATGTCCCGACCTCTCAATCCAGATGACGCTGCC 1200
Db 1141 TTTGGGAGGCTCCGAGCGCAACACATGATGTCCCGACCTCTCAATCCAGATGACGCTGCC 1200
QY 1201 AATCCCTGTGTAGGCTGTGAGTGTGCTCATCATCGCATCTTCTGGACAGCGGGAGCT 1260
Db 1201 AATCCCTGTGTAGGCTGTGAGTGTGCTCATCATCGCATCTTCTGGACAGCGGGAGCT 1260
QY 1261 GACTGCTCTGAGCAACCGAGAGCGCATCTCTGCGCAAGGATCTGCGGGGCGGC 1320
|||||

Db 1261 GACTGCTCTGAGCAACCGAGAGCGCATCTCTGCGCAAGGATCTGCGGGCGGC 1320
QY 1321 AGCTACACCTGTAGGCGAGAGTGGAGCTGGCTTTTGGCGTGGGCTCCAGCGCTGTCT 1380
Db 1321 AGCTACACCTGTAGGCGAGAGTGGAGCTGGCTTTTGGCGTGGGCTCCAGCGCTGTCT 1380
QY 1381 TACATGCACTACTGACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGATGTG 1440
Db 1381 TACATGCACTACTGACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGATGTG 1440
QY 1441 CAGACCGGCACTTCCCTGGGCGCATGGGAGCAGCAGCTGTGGGAGAGGACATCTGCTC 1500
Db 1441 CAGACCGGCACTTCCCTGGGCGCATGGGAGCAGCAGCTGTGGGAGAGGACATCTGCTC 1500
QY 1501 AAGGGGCGCTGTGGGAGAGACAACTCAACAGCAAGGAGTGTGATGTTCTGGGCGC 1560
Db 1501 AAGGGGCGCTGTGGGAGAGACAACTCAACAGCAAGGAGTGTGATGTTCTGGGCGC 1560
QY 1561 AATGGGATCCCTATGTGGCCCTGTGCGCACATGTGGTGGGGGCTGACCTGCGCAGG 1620
Db 1561 AATGGGATCCCTATGTGGCCCTGTGCGCACATGTGGTGGGGGCTGACCTGCGCAGG 1620
QY 1621 AGGAGTGTACCAACCCCAACCCCTGCAAGCGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
Db 1621 AGGAGTGTACCAACCCCAACCCCTGCAAGCGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
QY 1681 AATACCGATCTCTGCAATCTGGAGCGCTGCCACCTCAGCTCCGGAAGAAGCTTCGGG 1740
Db 1681 AATACCGATCTCTGCAATCTGGAGCGCTGCCACCTCAGCTCCGGAAGAAGCTTCGGG 1740
QY 1741 GAGGAGCACTGTGAGGCTTTCAAGGCTTACAAACAGCAGCAGCAGCGCTCATCTGCGC 1800
Db 1741 GAGGAGCACTGTGAGGCTTTCAAGGCTTACAAACAGCAGCAGCAGCGCTCATCTGCGC 1800
QY 1801 GTGGAGTGTGGTCCCAAGTACTCCGCGGTGTCCCGGAGCAAGTGAAGTCACTGCTGCG 1860
Db 1801 GTGGAGTGTGGTCCCAAGTACTCCGCGGTGTCCCGGAGCAAGTGAAGTCACTGCTGCG 1860
QY 1861 CGAGCCATGTGCACTGTCTATGTGTGTCGAGCCCAAGAGTGTGAGGCGCACGCTG 1920
Db 1861 CGAGCCATGTGCACTGTCTATGTGTGTCGAGCCCAAGAGTGTGAGGCGCACGCTG 1920
QY 1921 TGTCTCCCTGACTCCACCTCGCTGTGTGTCAGGCAAGTGTATCAAGGCTGTGAT 1980
Db 1921 TGTCTCCCTGACTCCACCTCGCTGTGTGTCAGGCAAGTGTATCAAGGCTGTGAT 1980
QY 1981 GGAACCTGTGGCTCCAGAGAGATTTGCACAGTGTGGGCTGTGTGGGGAGACATTAAG 2040
Db 1981 GGAACCTGTGGCTCCAGAGAGATTTGCACAGTGTGGGCTGTGTGGGGAGACATTAAG 2040
QY 2041 AGCTGCAAGAGTGAAGTGTGACTTTCATGTCGAGCCCAAGCCATGATGCTACATTTCTGTG 2100
Db 2041 AGCTGCAAGAGTGAAGTGTGACTTTCATGTCGAGCCCAAGCCATGATGCTACATTTCTGTG 2100
QY 2101 GGCATCCCGGAGGGGCTTCAGCATCGCATCCGCGAGCGGGTTAAAGGCGCTGATC 2160
Db 2101 GGCATCCCGGAGGGGCTTCAGCATCGCATCCGCGAGCGGGTTAAAGGCGCTGATC 2160
QY 2161 GGGGATGACAACTACTGTGCTGTGAAGAACAAGCAAGGCAAGTACTCTCAAGCGGAT 2220
Db 2161 GGGGATGACAACTACTGTGCTGTGAAGAACAAGCAAGGCAAGTACTCTCAAGCGGAT 2220
QY 2221 TTTGTGTGTGTGGGCGGTGAGCGGAGCTGTGTGTAAGGGGAGTCTCTCGGTACAGC 2280
Db 2221 TTTGTGTGTGTGGGCGGTGAGCGGAGCTGTGTGTAAGGGGAGTCTCTCGGTACAGC 2280
QY 2281 GGCAGGGGACAGCGGTGAGAGCTGTGAGGCTTCCCGGCCATCTGTGAGGCGCTGAC 2340
Db 2281 GGCAGGGGACAGCGGTGAGAGCTGTGAGGCTTCCCGGCCATCTGTGAGGCGCTGAC 2340
QY 2341 GTGAGAGTCTCTCGGTGGGAGAGTGAACACCGCGCGGCTGCTTCTATCTG 2400
Db 2341 GTGAGAGTCTCTCGGTGGGAGAGTGAACACCGCGCGGCTGCTTCTATCTG 2400
|||||

Query Match	Best Local Similarity	Score	DB	Length
Matches 2852; Conservative	100.0%;	2802;	8;	2853;
	0;	Pred. No. 0;	Mismatches 1;	Indels 0;
			Gaps 0;	
1	ATGCTCTCTGGGCAATCTTAACCCGCGGTTTCGGGGGGAACCGCTGAGGCTCTGAG	60		
1	ATGCTCTCTGGGCAATCTTAACCCGCGGTTTCGGGGGGAACCGCTGAGGCTCTGAG	60		
61	CCAGACCGGGAGGTACTGCTCCCATCCGACTGAGACCCGAGACATTAACGGCGCGCTAC	120		
61	CCAGACCGGGAGGTACTGCTCCCATCCGACTGAGACCCGAGACATTAACGGCGCGCTAC	120		
121	TACTGGCGGGGTTCCGAGACATCCGGGGATCAAGGACTATTTTACATCACACATT	180		
121	TACTGGCGGGGTTCCGAGACATCCGGGGATCAAGGACTATTTTACATCACACATT	180		
181	CAGGAGGATTTTACTTAACCTTAACCTGAGACGCGGATGCTCACTTTGGCTCCGCTTTC	240		
181	CAGGAGGATTTTACTTAACCTTAACCTGAGACGCGGATGCTCACTTTGGCTCCGCTTTC	240		
241	ACTGAGCATCTGAGGCGTCCCTCCAGAGGGGCTACCGGGGGCTCTTCAGACCTGCGACG	300		
241	ACTGAGCATCTGAGGCGTCCCTCCAGAGGGGCTACCGGGGGCTCTTCAGACCTGCGACG	300		
301	TGCTTCTATTTCTGGGGAGCTGAACGCGGAGCTGCTGCTGCTGAGGCTGAG	360		
301	TGCTTCTATTTCTGGGGAGCTGAACGCGGAGCTGCTGCTGCTGAGGCTGAG	360		
361	GGGGGGGCTCCGCGAGGCTTTGGCTAACCGAGCGCGGAGTATGCTATTAGCGCGTGCC	420		
361	GGGGGGGCTCCGCGAGGCTTTGGCTAACCGAGCGCGGAGTATGCTATTAGCGCGTGCC	420		
421	AATGCTAGCGCGCGCGCGCGCGAGCGCAACAGCCAGGGCGACACCTTCCAGTCCGCG	480		
421	AATGCTAGCGCGCGCGCGCGCGAGCGCAACAGCCAGGGCGACACCTTCCAGTCCGCG	480		
481	GGTGTTTCGGGGGCTCTTCGAGACCCACCTCTCGTGGGGGTGGCGTCCGGGGCTGG	540		
481	GGTGTTTCGGGGGCTCTTCGAGACCCACCTCTCGTGGGGGTGGCGTCCGGGGCTGG	540		
541	AACCCGCGCATCTTAACGGGGCTTGAACCCCTTCAAGCCGCGGGCGGGGCTTCGGGGAG	600		
541	AACCCGCGCATCTTAACGGGGCTTGAACCCCTTCAAGCCGCGGGCGGGGCTTCGGGGAG	600		
601	AGTGTAGCGCGCGAGGAGTGGGGGGCGCGCAAGGCTTCTGCTATACCCGGGATGAG	660		
601	AGTGTAGCGCGCGAGGAGTGGGGGGCGCGCAAGGCTTCTGCTATACCCGGGATGAG	660		
661	GAGAGCGTGTGGTTCGCGAGAGAGTCAATGTGCAAGTTTCCAGGGCGGAGCTGAGACAT	720		
661	GAGAGCGTGTGGTTCGCGAGAGAGTCAATGTGCAAGTTTCCAGGGCGGAGCTGAGACAT	720		
721	TATCTGCTGAGCGTGTGGGCAACGGCGCGGACTCTTAACGCCCATCCAGCATCTCTAAC	780		
721	TATCTGCTGAGCGTGTGGGCAACGGCGCGGACTCTCTTAACGCCCATCCAGCATCTCTAAC	780		
781	CCCATCAACATCTGTTGTGTCAAGTGTCTTCAAGATCGTACCTCCGGGGCCAG	840		
781	CCCATCAACATCTGTTGTGTCAAGTGTCTTCAAGATCGTACCTCCGGGGCCAG	840		
841	GTCACCGGCAATGCGGCCCTGACGCTGCGCAATCTTGTGCTTGGCAGAGAACCTGAC	900		
841	GTCACCGGCAATGCGGCCCTGACGCTGCGCAATCTTGTGCTTGGCAGAGAACCTGAC	900		

QY 901 AAAGTAGACAAAGACCCCGAGTACTGGACACTGCCATCTCTTCAACAGGACAGAC 960
 Db 901 AAAGTAGACAAAGACCCCGAGTACTGGACACTGCCATCTCTTCAACAGGACAGAC 960
 QY 961 CTGTGTGGAGCCACACCTGTGACACCCCTGGGATGGCTGATGTGATCATGTGTGAC 1020
 Db 961 CTGTGTGGAGCCACACCTGTGACACCCCTGGGATGGCTGATGTGATCATGTGTGAC 1020
 QY 1021 CCCAAGGAAGCTGCTGTGTGATTTGAGAGAGATGGGTTCCATCAGCCTTACACTTGGC 1080
 Db 1021 CCCAAGGAAGCTGCTGTGTGATTTGAGAGAGATGGGTTCCATCAGCCTTACACTTGGC 1080
 QY 1081 CACGAGCTGGGACCAAGCTGTCAACATGCCCATGACATGTGAAGTCTTGAAGAGTGTG 1140
 Db 1081 CACGAGCTGGGACCAAGCTGTCAACATGCCCATGACATGTGAAGTCTTGAAGAGTGTG 1140
 QY 1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGACCTCATCCAGATCSACCTGGC 1200
 Db 1141 TTTGGGAAGCTCCGAGCCAAACACATGATGTCCCGACCTCATCCAGATCSACCTGGC 1200
 QY 1201 AACCCGTGACAGCTGCTGATGGCTGCATCATCAGCACTTCTGGACAGCGGACAGGT 1260
 Db 1201 AACCCGTGACAGCTGCTGATGGCTGCATCATCAGCACTTCTGGACAGCGGACAGGT 1260
 QY 1261 GACTGCTCTGTGACCAACACCAAGCCATCTCCCTGGCCGAGATCTGCCGGAGGCGC 1320
 Db 1261 GACTGCTCTGTGACCAACACCAAGCCATCTCCCTGGCCGAGATCTGCCGGAGGCGC 1320
 QY 1321 AGCTACACCTTGAGCCAGAGCTGTGGCTTTTGGCTGGGCTTCCAGCCCTGTCT 1380
 Db 1321 AGCTACACCTTGAGCCAGAGCTGTGGCTTTTGGCTGGGCTTCCAGCCCTGTCT 1380
 QY 1381 TACATGAGTACTGCTACCAAGCTGTGTGACACCGGAGGACCAAGGACAGATGTGTG 1440
 Db 1381 TACATGAGTACTGCTACCAAGCTGTGTGACACCGGAGGACCAAGGACAGATGTGTG 1440
 QY 1441 CAGACCCGCACTTCCCTGGGCGGATGACACAGCTGTGGGAGGAGCAAGCTGTGCTC 1500
 Db 1441 CAGACCCGCACTTCCCTGGGCGGATGACACAGCTGTGGGAGGAGCAAGCTGTGCTC 1500
 QY 1501 AAAAGGGGCTGCGTGGAGAGACACACCTGACAAAGACAGAGGATGTCTCTGGGCGC 1560
 Db 1501 AAAAGGGGCTGCGTGGAGAGACACACCTGACAAAGACAGAGGATGTCTCTGGGCGC 1560
 QY 1561 AAATGGGATCCCTATGGGCGCTGCTCGGACATGTGTGGGCGGTGACAGTGTGCGCAG 1620
 Db 1561 AAATGGGATCCCTATGGGCGCTGCTCGGACATGTGTGGGCGGTGACAGTGTGCGCAG 1620
 QY 1621 AGGCAATGACCAACCCACCTGCGCAAGGGGGCAAGTACTGCGAGGAGTGTGAGG 1680
 Db 1621 AGGCAATGACCAACCCACCTGCGCAAGGGGGCAAGTACTGCGAGGAGTGTGAGG 1680
 QY 1681 AAATACGATCTGCAATCTGAGAGCCCTGCGCCAGCTCAGCTCCGAAAGAGCTTCCG 1740
 Db 1681 AAATACGATCTGCAATCTGAGAGCCCTGCGCCAGCTCAGCTCCGAAAGAGCTTCCG 1740
 QY 1741 GAGGAGCAAGTGTGAGGCTTTCACAGGCTACAAACACAGCAACCGGCTCAGCTTGGC 1800
 Db 1741 GAGGAGCAAGTGTGAGGCTTTCACAGGCTACAAACACAGCAACCGGCTCAGCTTGGC 1800
 QY 1801 GTGGCATGGGTGGCCAAAGTACTCCGGGCTGTCTCCCGGGCAAGTCAAGCTCATCTGC 1860
 Db 1801 GTGGCATGGGTGGCCAAAGTACTCCGGGCTGTCTCCCGGGCAAGTCAAGCTCATCTGC 1860
 QY 1861 CGAGCAATGGCACTGGCTACTTCTATGTGTGACCAAGGAGTGTGAGGAGGAGGAGCTG 1920
 Db 1861 CGAGCAATGGCACTGGCTACTTCTATGTGTGACCAAGGAGTGTGAGGAGGAGGAGCTG 1920
 QY 1921 TGCTCTCTGACTCCACTCCGCTGTGTGTGCCAAGGAGTGCATCAAGGCTGGCTGTGAT 1980
 Db 1921 TGCTCTCTGACTCCACTCCGCTGTGTGTGCCAAGGAGTGCATCAAGGCTGGCTGTGAT 1980

QY 1981 GGGAACTGTGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAAATAG 2040
 Db 1981 GGGAACTGTGGCTCCAAAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAAATAG 2040
 QY 2041 AGCTGCAAGAGTGTGACTGTGACTCTTCAACCAAGCCATGATGTGCTACATTTGTGGT 2100
 Db 2041 AGCTGCAAGAGTGTGACTGTGACTCTTCAACCAAGCCATGATGTGCTACATTTGTGGT 2100
 QY 2101 GGCATCCCCGAGGGGCTTCAAGCATGACATCCGCGAGCCGCTTCAAGGCGCTATC 2160
 Db 2101 GGCATCCCCGAGGGGCTTCAAGCATGACATCCGCGAGCCGCTTCAAGGCGCTATC 2160
 QY 2161 GGGGATGACAACTACTGGCTCTGTAAGAACCCAGGCAAGTACTGCTCAACGGGCTAT 2220
 Db 2161 GGGGATGACAACTACTGGCTCTGTAAGAACCCAGGCAAGTACTGCTCAACGGGCTAT 2220
 QY 2221 TTGCTGTGTGTGGGCTGTGAGCGGAGCTGTGTGTAAGGGGAGTGTGCTGAGTACAGC 2280
 Db 2221 TTGCTGTGTGTGGGCTGTGAGCGGAGCTGTGTGTAAGGGGAGTGTGCTGAGTACAGC 2280
 QY 2281 GGCAGGGGACAGCGGTGGAGAGCTGTGAGAGCTTCCCGGCCATCTGAGCCGCTGAC 2340
 Db 2281 GGCAGGGGACAGCGGTGGAGAGCTGTGAGAGCTTCCCGGCCATCTGAGCCGCTGAC 2340
 QY 2341 GTGAGTCTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCCGCTACTCTTATCTG 2400
 Db 2341 GTGAGTCTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCCGCTACTCTTATCTG 2400
 QY 2401 CCCAAAGAGCCTGGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGAGCCCTGTGTC 2460
 Db 2401 CCCAAAGAGCCTGGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGAGCCCTGTGTC 2460
 QY 2461 TTGCAACAAAGGCTCCTAGCCTCTCAACCAAGTGTGAGCAGCGGACAGAGCCCTT 2520
 Db 2461 TTGCAACAAAGGCTCCTAGCCTCTCAACCAAGTGTGAGCAGCGGACAGAGCCCTT 2520
 QY 2521 GCACGCTGGGTGGTGGAGAGCTGGGGGCGGCTCTCCGGAGCTGTGGGAGTGGCTGAG 2580
 Db 2521 GCACGCTGGGTGGTGGAGAGCTGGGGGCGGCTCTCCGGAGCTGTGGGAGTGGCTGAG 2580
 QY 2581 AAGCGGCGGTGAGTGTGTGCGGCGCTCCGCGGGAGCCAGCGTCCCTGCTGATGCA 2640
 Db 2581 AAGCGGCGGTGAGTGTGTGCGGCGCTCCGCGGGAGCCAGCGTCCCTGCTGATGCA 2640
 QY 2641 GCCCATGGCCCGGTGGAGACACAGCTGTGCGGGAGGCCCTGCTCCACTGGAGCTCAGC 2700
 Db 2641 GCCCATGGCCCGGTGGAGACACAGCTGTGCGGGAGGCCCTGCTCCACTGGAGCTCAGC 2700
 QY 2701 GCTGTGCTACCTGTGCTCAAGAGCTGGGGCGGGGATTTCAAGAGGCGTCAAGTGT 2760
 Db 2701 GCTGTGCTACCTGTGCTCAAGAGCTGGGGCGGGGATTTCAAGAGGCGTCAAGTGT 2760
 QY 2761 GTGGGCGACAGGAGCGGCTGTGCGCCGCGGAGCCAGTGCATCTTACACGCAAGCCGAG 2820
 Db 2761 GTGGGCGACAGGAGCGGCTGTGCGCCGCGGAGCCAGTGCATCTTACACGCAAGCCGAG 2820
 QY 2821 GAGCTGAGCTTGTGGCTGTGAGGCGGCTGTCA 2853
 Db 2821 GAGCTGAGCTTGTGGCTGTGAGGCGGCTGTCA 2853

RESULT 3
 US-10-170-235-10538
 ; Sequence 10538, Application US/10170235
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU
 ; FILE REFERENCE: CL001380
 ; CURRENT APPLICATION NUMBER: US/10/170,235
 ; NUMBER OF SEQ ID NOS: 42514
 ; SEQ ID NO 10538

QY	961	CTGTGTGGAGCCACACACTGTGTGACACACCTTGGCATGTGGGTACCATGTGTAC	1020
Db	961	CTGTGTGGAGCCACACACTGTGTGACACACCTTGGCATGTGGGTACCATGTGTAC	1020
QY	1021	CCCAAGAGAAAGCTGTCTGTATGTAGAGACGATGGGTTTCATCAACCTTCAACACTCC	1080
Db	1021	CCCAAGAGAAAGCTGTCTGTATGTAGAGACGATGGGTTTCATCAACCTTCAACACTCC	1080
QY	1081	CACGAGCTGGGCGACAGTGTTCACATGCCCCATGACAAATGTGAAAGTCTGTAGAGAGTG	1140
Db	1081	CACGAGCTGGGCGACAGTGTTCACATGCCCCATGACAAATGTGAAAGTCTGTAGAGAGTG	1140
QY	1141	TTTTGGGAAGCTCCGAGGCCAACACATATATGTCCCGACCTCATCCAGATGCACCTGTCC	1200
Db	1141	TTTTGGGAAGCTCCGAGGCCAACACATATATGTCCCGACCTCATCCAGATGCACCTGTCC	1200
QY	1201	AACCCCTGTAGAGCTGAGTGGTGCACATATACGCACTTCTGTGACAGCGGACGCT	1260
Db	1201	AACCCCTGTAGAGCTGAGTGGTGCACATATACGCACTTCTGTGACAGCGGACGCT	1260
QY	1261	GACTGCTCTCTGGACCAACCCAGACGCCATCTCCCTCCCGAGAGTCTGCCGGGCGCC	1320
Db	1261	GACTGCTCTCTGGACCAACCCAGACGCCATCTCCCTCCCGAGAGTCTGCCGGGCGCC	1320
QY	1321	AGCTACACCTGTGAGCCAGAGTGGAGAGCTTGTGGGTGGGGTCCAAAGCTGTGCT	1380
Db	1321	AGCTACACCTGTGAGCCAGAGTGGAGAGCTTGTGGGTGGGGTCCAAAGCTGTGCT	1380
QY	1381	TACATGACGATCTGACCAACAGCTGTGTGTGACCCGGGAAGGCCMAAGGACAGATGGTGTGC	1440
Db	1381	TACATGACGATCTGACCAACAGCTGTGTGTGACCCGGGAAGGCCMAAGGACAGATGGTGTGC	1440
QY	1441	CAGACCCGACCTTCCCTTGGGGCGAATGGACAGTGTGGAGAGGCAACCTCTGCTC	1500
Db	1441	CAGACCCGACCTTCCCTTGGGGCGAATGGACAGTGTGGAGAGGCAACCTCTGCTC	1500
QY	1501	AAAGGGGCTGCGTGGAGAGACAAACCTCAACAAACAGAGGTGGATGCTTCTGGGCC	1560
Db	1501	AAAGGGGCTGCGTGGAGAGACAAACCTCAACAAACAGAGGTGGATGCTTCTGGGCC	1560
QY	1561	AAATGGGATTCCTATGTGGCCCCCTGCTGGCGCACATGTGGTGGGGGCTGACCTGCCAAG	1620
Db	1561	AAATGGGATTCCTATGTGGCCCCCTGCTGGCGCACATGTGGTGGGGGCTGACCTGCCAAG	1620
QY	1621	AGGCAAGTGCACAAACCCACCCCTGCGCAACAGGGGGCAATATCTGGAGGGAGTGAAGGTG	1680
Db	1621	AGGCAAGTGCACAAACCCACCCCTGCGCAACAGGGGGCAATATCTGGAGGGAGTGAAGGTG	1680
QY	1681	AAATATCGATCTGTCAATCTTGGAGGCCCTGCCCCAGTCAAGCTCCGGAAGAAGCTTCCAG	1740
Db	1681	AAATATCGATCTGTCAATCTTGGAGGCCCTGCCCCAGTCAAGCTCCGGAAGAAGCTTCCAG	1740
QY	1741	GAGGAGAGTGTGAGGCTTTCACAGGCTCAACACAGAGACCAACCGGCTACCTGTGCC	1800
Db	1741	GAGGAGAGTGTGAGGCTTTCACAGGCTCAACACAGAGACCAACCGGCTACCTGTGCC	1800
QY	1801	GTGCGATAGGTGCCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTGCACATCATCTGC	1860
Db	1801	GTGCGATAGGTGCCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTGCACATCATCTGC	1860
QY	1861	CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGGACCCAAAGGTGTGACAGGCAAGCTG	1920
Db	1861	CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGGACCCAAAGGTGTGACAGGCAAGCTG	1920
QY	1921	TGCTCTCTGACTTCACACTCCGTCTGTGTCCAAGGCAAGTGCATCAAGGCTGTGTGAT	1980
Db	1921	TGCTCTCTGACTTCACACTCCGTCTGTGTCCAAGGCAAGTGCATCAAGGCTGTGTGAT	1980
QY	1981	GGGAACTGGGGCTTCCAAAGAGATTTCGCAAGTGTGGGGGTGTGGGGGAGACAAATTAAG	2040
Db	1981	GGGAACTGGGGCTTCCAAAGAGATTTCGCAAGTGTGGGGGTGTGGGGGAGACAAATTAAG	2040
QY	2041	AGCTGCAAAGAGTGTGACTGAGCTCTTTCACAAAGGCCATGAGTGCATATTCCTGTGGGTG	2100

|||||
Db 961 CTGTGTGGAGCCACACCTGTGACACCCCTGGGCACTGATGTGGTACATGTGTGAC 1020
|||
OY 1021 CCCAAGAGAAAGCTCTCTGTATGTAGAGAGATGGGCTTCCATCAGCTTCACCACTGCC 1080
|||
Db 1021 CCCAAGAGAAAGCTCTCTGTATGTAGAGAGATGGGCTTCCATCAGCTTCACCACTGCC 1080
|||
OY 1081 CACGAGCTGGGGCCAGTGTTCACATGCCCATGACAAATGTGAAGTCTGTAGAGAGTGG 1140
|||
Db 1081 CACGAGCTGGGGCCAGTGTTCACATGCCCATGACAAATGTGAAGTCTGTAGAGAGTGG 1140
|||
OY 1141 TTGTGGAGAGCTCCGAGCAACACATGATGTCCCGACCTCTCATCCAGATGACCGTGGC 1200
|||
Db 1141 TTGTGGAGAGCTCCGAGCAACACATGATGTCCCGACCTCTCATCCAGATGACCGTGGC 1200
|||
OY 1201 AACCCCTGGTCTGAGCTGATGCTGATCATACGAGCTTCTGTGACAGAGGGGACGGT 1260
|||
Db 1201 AACCCCTGGTCTGAGCTGATGCTGATCATACGAGCTTCTGTGACAGAGGGGACGGT 1260
|||
OY 1261 GACTGCTCTGTGAGCAACCCAGCAAGCCCATCTCCCTGGCCGAGGATCTGCCGGGCGCC 1320
|||
Db 1261 GACTGCTCTGTGAGCAACCCAGCAAGCCCATCTCCCTGGCCGAGGATCTGCCGGGCGCC 1320
|||
OY 1321 AGCTACACCTGAGCCAGAGCTGCGAGCTGTGCTTTGGGCTGGCTCCAAAGCCCTGTCT 1380
|||
Db 1321 AGCTACACCTGAGCCAGAGCTGCGAGCTGTGCTTTGGGCTGGCTCCAAAGCCCTGTCT 1380
|||
OY 1381 TACATGAGTACTGACACCAAGCTGTGTGACAGGGGCAAGGAGACAGATGTGTGTC 1440
|||
Db 1381 TACATGAGTACTGACACCAAGCTGTGTGACAGGGGCAAGGAGACAGATGTGTGTC 1440
|||
OY 1441 CAGACCCGCACTTCCCTGTGGGCGATGTGACACAGCTGTGGCCAGGCAACCTCTGCTC 1500
|||
Db 1441 CAGACCCGCACTTCCCTGTGGGCGATGTGACACAGCTGTGGCCAGGCAACCTCTGCTC 1500
|||
OY 1501 AAAGGGGCTGGTGGAGAGACACAACTCTCAACACAGAGGTGATGCTTCTGTGGCC 1560
|||
Db 1501 AAAGGGGCTGGTGGAGAGACACAACTCTCAACACAGAGGTGATGCTTCTGTGGCC 1560
|||
OY 1561 AAATGGGATCCCTATGTGGCCCTGTGCGGACATGTGTGGGGGCGCTGACAGCTGGCCAG 1620
|||
Db 1561 AAATGGGATCCCTATGTGGCCCTGTGCGGACATGTGTGGGGGCGCTGACAGCTGGCCAG 1620
|||
OY 1621 AGCAGTGTACCAACCCCAACCCCTGTGCAAGGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
|||
Db 1621 AGCAGTGTACCAACCCCAACCCCTGTGCAAGGGGGCAAGTACTGCGAGGAGTGAAGGTG 1680
|||
OY 1681 AAATACGATCCCTGCAATGTGAGCGCTGCGCCAGCTCAGGCTCCGGAAGAGCTTCCGG 1740
|||
Db 1681 AAATACGATCCCTGCAATGTGAGCGCTGCGCCAGCTCAGGCTCCGGAAGAGCTTCCGG 1740
|||
OY 1741 GAGGAGCAGTGTGAGGCTTTCACAGGCTACACACAGACAGACCGGCTCACTCTCGCC 1800
|||
Db 1741 GAGGAGCAGTGTGAGGCTTTCACAGGCTACACACAGACAGACCGGCTCACTCTCGCC 1800
|||
OY 1801 GTGGGATGGGTGCCCAAGTACTCCGGGCTGTCTCCCGGGACAAGTCAATCTGTC 1860
|||
Db 1801 GTGGGATGGGTGCCCAAGTACTCCGGGCTGTCTCCCGGGACAAGTCAATCTGTC 1860
|||
OY 1861 CGAGCCATGGGAGCTGCTACTTCTATGTGCTGAGCAACCAAGCTGTGAGACGGACGCTG 1920
|||
Db 1861 CGAGCCATGGGAGCTGCTACTTCTATGTGCTGAGCAACCAAGCTGTGAGACGGACGCTG 1920
|||
OY 1921 TGTCTCTGACTCTCACTCTGCTGTGTGCAAGGCAAGTGAATGAAGGCTGGCTGTGAT 1980
|||
Db 1921 TGTCTCTGACTCTCACTCTGCTGTGTGCAAGGCAAGTGAATGAAGGCTGGCTGTGAT 1980
|||
OY 1981 GGGAACTGGGCTCCCAAGAGAGATTCGACAAGTGTGGGCTGTGGGGGGAGACAATAAG 2040
|||
Db 1981 GGGAACTGGGCTCCCAAGAGAGATTCGACAAGTGTGGGCTGTGGGGGGAGACAATAAG 2040
|||
OY 2041 AGCTGCAAGAGGTGACTGTGCTTTCACCAAGCCCATGCAATGCTTCTGTG 2100
|||

Db 2041 AGCTGCAAGAGGTGACTGTGCTTTCACCAAGCCCATGCAATGCTTCTGTG 2100
|||
OY 2101 GCCATCCCCGAGGCGCTCAAGCATCGACATCGCCAGCGGCTTACAAAGGCTGTATC 2160
|||
Db 2101 GCCATCCCCGAGGCGCTCAAGCATCGACATCGCCAGCGGCTTACAAAGGCTGTATC 2160
|||
OY 2161 GGGGATGACACTTACTGCTGTGAAGAAAGCCCAAGGCAAGTACTCTCTCAAGCGCAT 2220
|||
Db 2161 GGGGATGACACTTACTGCTGTGAAGAAAGCCCAAGGCAAGTACTCTCTCAAGCGCAT 2220
|||
OY 2221 TTCGTGTGTGCGCGGTGAGCGGAGCTGTGTGTGAAGGGGAGTCTCTCGGTACAGC 2280
|||
Db 2221 TTCGTGTGTGCGCGGTGAGCGGAGCTGTGTGTGAAGGGGAGTCTCTCGGTACAGC 2280
|||
OY 2281 GGCAGGGGACAGCGGTGAGAGCTTGCAGAGCTTCCCGGCCCATCTGTGAGCCGCTGAC 2340
|||
Db 2281 GGCAGGGGACAGCGGTGAGAGCTTGCAGAGCTTCCCGGCCCATCTGTGAGCCGCTGAC 2340
|||
OY 2341 GTGGAGGTCTCTCCGTGGGGAAGTACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
|||
Db 2341 GTGGAGGTCTCTCCGTGGGGAAGTACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
|||
OY 2401 CCCAAAGAGCTCTGGGAGAGACAACTCTCTCATCCCAAGAGACCCCGGGAGCCCTGTGTC 2460
|||
Db 2401 CCCAAAGAGCTCTGGGAGAGACAACTCTCTCATCCCAAGAGACCCCGGGAGCCCTGTGTC 2460
|||
OY 2461 TTGCACAAACAGGTCTCTCAAGCTCTTCAACCAAGTGTGAGAGCCGGAGCAGAGGCCCT 2520
|||
Db 2461 TTGCACAAACAGGTCTCTCAAGCTCTTCAACCAAGTGTGAGAGCCGGAGCAGAGGCCCT 2520
|||
OY 2521 GCACGCTGGTGGTGGGAGCTGGGGGGCGTGTCTCCGGAGCTGGGGAGTGGGCTGTGAG 2580
|||
Db 2521 GCACGCTGGTGGTGGGAGCTGGGGGGCGTGTCTCCGGAGCTGGGGAGTGGGCTGTGAG 2580
|||
OY 2581 AAGCGGGCGGTGAGTGTGCGGGCTCCGCGGGAGCGCAGCGTCCCTGTGTATGCA 2640
|||
Db 2581 AAGCGGGCGGTGAGTGTGCGGGCTCCGCGGGAGCGCAGCGTCCCTGTGTATGCA 2640
|||
OY 2641 GCGCATGCGGCGGTGAGACACAAAGCTTGGGGGAGCCCTCCCTGAGGAGCTCAAGCT 2700
|||
Db 2641 GCGCATGCGGCGGTGAGACACAAAGCTTGGGGGAGCCCTCCCTGAGGAGCTCAAGCT 2700
|||
OY 2701 GCTTGTACACCTGTCTCAAGAGCTGCGCGGGGATTTGAGAGCGCTCACTCAAGTGT 2760
|||
Db 2701 GCTTGTACACCTGTCTCAAGAGCTGCGCGGGGATTTGAGAGCGCTCACTCAAGTGT 2760
|||
OY 2761 GTGGGCAACGAGAGCGGCTGTGCGCGGGAGCAGTGTGCAACCGGACCGCCAG 2820
|||
Db 2761 GTGGGCAACGAGAGCGGCTGTGCGCGGGAGCAGTGTGCAACCGGACCGCCAG 2820
|||
OY 2821 GAGCTGAGCTTCTGCTCTGAGGCGCGGTGCA 2853
|||
Db 2821 GAGCTGAGCTTCTGCTCTGAGGCGCGGTGCA 2853
|||

RESULT 5
US-60-453-135-7278
Sequence 7278, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARILLI, Michele
APPLICANT: IAKUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CI001456
CURRENT APPLICATION NUMBER: US/60/453,135
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7278
LENGTH: 2853
TYPE: DNA
ORGANISM: Homo sapiens
US-60-453-135-7278

QY	1021	CCCAAGGAAGCTGCTCTGTCAATTGAGGACCAATGGGGCTTCCATCAAGCTTACCACTGGC	1080
Db	1021	CCCAAGGAAGCTGCTCTGTCAATTGAGGACCAATGGGGCTTCCATCAAGCTTACCACTGGC	1080
QY	1081	CACAGCTGGGGCCACGGTTCAACATGAGCCCATGTGACAAATGTGAAAGTCTGTGAGAGAGTG	1140
Db	1081	CACAGCTGGGGCCACGGTTCAACATGAGCCCATGTGACAAATGTGAAAGTCTGTGAGAGAGTG	1140
QY	1201	AAACCCGTGGTACGCTGTGAGTGGTGGCCATCATCACGAGCTTCGTGGAGAGGGCCACGGT	1260
Db	1201	AAACCCGTGGTACGCTGTGAGTGGTGGCCATCATCACGAGCTTCGTGGAGAGGGCCACGGT	1260
QY	1261	GACTGCTCCTGGAGCAACCCAGCAAGCCCATCTCCCTGGAGGATCTGGCGGGCCG	1320
Db	1261	GACTGCTCCTGGAGCAACCCAGCAAGCCCATCTCCCTGGAGGATCTGGCGGGCCG	1320
QY	1321	AGCTACACCTGAGCCAGCAGTGGCCGAGTGGCTTTTGGCGTGGGCTCCAAACCTGTGCT	1380
Db	1321	AGCTACACCTGAGCCAGCAGTGGCCGAGTGGCTTTTGGCGTGGGCTCCAAACCTGTGCT	1380
QY	1381	TACATGAGTACTGACACCAAGCTGTGTGTACACCGGGAAAGGCCAAGGACAGATGTGTGC	1440
Db	1381	TACATGAGTACTGACACCAAGCTGTGTGTACACCGGGAAAGGCCAAGGACAGATGTGTGC	1440
QY	1441	CAGACCCGCACTCTCCCTGTGGGCCCATGTGGCACCACTGTGGCGAGGGCAACCTGTGCTC	1500
Db	1441	CAGACCCGCACTCTCCCTGTGGGCCCATGTGGCACCACTGTGGCGAGGGCAACCTGTGCTC	1500
QY	1501	AAAGGGGCGTGGCGTGGAGAGACACCACTGTCAACAGCAGCAGGGTGTGATGTCTCTGGGCC	1560
Db	1501	AAAGGGGCGTGGCGTGGAGAGACACCACTGTCAACAGCAGCAGGGTGTGATGTCTCTGGGCC	1560
QY	1561	AAATGGGATCCCTATGTGGCCCTGTGCTGGCGACATGTGTGGGGGGCTGACCTGGCCAGG	1620
Db	1561	AAATGGGATCCCTATGTGGCCCTGTGCTGGCGACATGTGTGGGGGGCTGACCTGGCCAGG	1620
QY	1621	AGGAGTGGACCAACCCCACTCCCTGGCAACGGGGGCAATATCTGGAGGAGTGAAGGTG	1680
Db	1621	AGGAGTGGACCAACCCCACTCCCTGGCAACGGGGGCAATATCTGGAGGAGTGAAGGTG	1680
QY	1741	GAGGAGCAGTGTGAGGCTTTCACAGGGCTGTCAACACAGCAGCAGCCGGCTACCTGTGGCC	1800
Db	1741	GAGGAGCAGTGTGAGGCTTTCACAGGGCTGTCAACACAGCAGCAGCCGGCTACCTGTGGCC	1800
QY	1801	GTGGCATGGGTGCCAAGTACTCGGCGTGTCTCCCGGGAGCAAGTCAAGCTCATCTGC	1860
Db	1801	GTGGCATGGGTGCCAAGTACTCGGCGTGTCTCCCGGGAGCAAGTCAAGCTCATCTGC	1860
QY	1861	CGAGCCAAATGGCACTGGCTACTCTTATGTGTGCTGGACCCCAAGGTGTGTGAGGACGCTG	1920
Db	1861	CGAGCCAAATGGCACTGGCTACTCTTATGTGTGCTGGACCCCAAGGTGTGTGAGGACGCTG	1920
QY	1921	TGCTCTCCTGTACTGACCTCCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1980
Db	1921	TGCTCTCCTGTACTGACCTCCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1980
QY	1981	GGGAACTGGGCTCCAGAAAGATTCGACAGTGTGGGGTGTGTGGGGAGACAAATTAAG	2040
Db	1981	GGGAACTGGGCTCCAGAAAGATTCGACAGTGTGGGGTGTGTGGGGAGACAAATTAAG	2040
QY	2041	AGCTCAAGAGGTGACTGGGACTCTTACCAACGCCCATCATGCTACAAATTTCTGGTGTG	2100
Db	2041	AGCTCAAGAGGTGACTGGGACTCTTACCAACGCCCATCATGCTACAAATTTCTGGTGTG	2100
QY	2101	GCCATTCGCCGAGGCGGCTCAAGCATTCGACATTCGCGAGCGCGTTACAAAGGCTGTATC	2160

Dp	1	ATGCTTGTGCTGGGCAATCTAAACCTTGACTTTGCGCGGGGGAAACCGGTGGAGGCTCTGAG	60
Qy	61	CCAGAGCGGAGAGTAAAGTGTGTCCATCCGAGCTGGAACCGGACATTAACGGCGCGCTTAC	120
Dp	61	CCAGAGCGGAGAGTAAAGTGTGTCCATCCGAGCTGGAACCGGACATTAACGGCGCGCTTAC	120
Qy	121	TACTGGCGGGGTCCCGAGAGACTCCGGGGATCAGGGACTCATTTTTCAGATCAGACATTT	180
Dp	121	TACTGGCGGGGTCCCGAGAGACTCCGGGGATCAGGGACTCATTTTTCAGATCAGACATTT	180
Qy	181	CAGAGAGACTTTTACCTACACTCAGCGCGGATCTCAGTTCTTGCGTCCCGCTTCTTC	240
Dp	181	CAGAGAGACTTTTACCTACACTCAGCGCGGATCTCAGTTCTTGCGTCCCGCTTCTTC	240
Qy	241	ACTGAGCATCTGGGGCTCCCCCTCCAGAGGGCTCACCGGGGGCTTTCAGACCTGCAGC	300
Dp	241	ACTGAGCATCTGGGGCTCCCCCTCCAGAGGGCTCACCGGGGGCTTTCAGACCTGCAGC	300
Qy	301	TGCTTCATTTCCTGGGGAGAGTAAAGCGCGGACACCTGTCGTGTGTGAGACCTGTGC	360
Dp	301	TGCTTCATTTCCTGGGGAGAGTAAAGCGCGGACACCTGTCGTGTGTGAGACCTGTGC	360
Qy	361	GGGGGGCTCCGCGGAGGCTTTGGCTACCGAGGGCGCGAGTATGATCATTAACCCGCTGCC	420
Dp	361	GGGGGGCTCCGCGGAGGCTTTGGCTACCGAGGGCGCGAGTATGATCATTAACCCGCTGCC	420
Qy	421	AATGCTAAGCGCGCGGGGGGCGGACGGAACAGCAGGGCGCACACTTCTTCAGCGCGCG	480
Dp	421	AATGCTAAGCGCGCGGGGGGCGGACGGAACAGCAGGGCGCACACTTCTTCAGCGCGCG	480
Qy	481	GGTGTTCGGGGCGGGGCTTCCGAGAGACCCCACTCTGTGCGAGGGGCGGCTTGCGCTGG	540
Dp	481	GGTGTTCGGGGCGGGGCTTCCGAGAGACCCCACTCTGTGCGAGGGGCGGCTTGCGCTGG	540
Qy	541	AACCCCGCACTCTACGAGGGGCTTGAGACCTTACAAAGCCGCGCGGGCGGGCTTCGAGAG	600
Dp	541	AACCCCGCACTCTACGAGGGGCTTGAGACCTTACAAAGCCGCGCGGGCGGGCTTCGAGAG	600
Qy	601	AGTGTAGCGCGCGCAGGCTGCGGGCGGGCAAAGGTTTGTGTCATCCGCGCTACGAG	660
Dp	601	AGTGTAGCGCGCGCAGGCTGCGGGCGGGCAAAGGTTTGTGTCATCCGCGCTACGAG	660
Qy	661	GAGAGCGTGGTGTGCGCGGAGAGTCAATGTCAAGTTTCAGCGCGCGGACCTGGAACAT	720
Dp	661	GAGAGCGTGGTGTGCGCGGAGAGTCAATGTCAAGTTTCAGCGCGCGGACCTGGAACAT	720
Qy	721	TATCTGCGAGCGGCTGGGGAACGGGGGCGGACCTACCGGCATCCGAGCATCTGCAC	780
Dp	721	TATCTGCGAGCGGCTGGGGAACGGGGGCGGACCTACCGGCATCCGAGCATCTGCAC	780
Qy	781	CCCATCAACATCGTTGTGTGAAGGTGCTCTCTTAGAGATGTGACCTCGCGCCAG	840
Dp	781	CCCATCAACATCGTTGTGTGAAGGTGCTCTCTTAGAGATGTGACCTCGCGCCAG	840
Qy	841	GTCACCGGGCAATGGGGCCCTGAGAGGCTGCGGACACTTCTGTGCTCTGGCAGAAAGCTGAAC	900
Dp	841	GTCACCGGGCAATGGGGCCCTGAGAGGCTGCGGACACTTCTGTGCTCTGGCAGAAAGCTGAAC	900
Qy	901	AAATGTGTACAAGCACCCCGAGTACTGGGACACTGCCATCTCTTACACAGGCGAGAC	960
Dp	901	AAATGTGTACAAGCACCCCGAGTACTGGGACACTGCCATCTCTTACACAGGCGAGAC	960
Qy	961	CTGTGTGAGCCACACACTGTGACACCCCTGGGCAATGGCTGATGTGGATCAATGTGTGAC	1020
Dp	961	CTGTGTGAGCCACACACTGTGACACCCCTGGGCAATGGCTGATGTGGATCAATGTGTGAC	1020
Qy	1021	CCCAAGGAAGCGCTGTGTCATTTAGAGAGAGTGGGCTTCACAGGCTTTCAGACACTGCC	1080
Dp	1021	CCCAAGGAAGCGCTGTGTCATTTAGAGAGAGTGGGCTTCACAGGCTTTCAGACACTGCC	1080
Qy	1081	CACGAGCTGGGCACTGTTCACAATGACCCCAATGACAAATGTAAAGTCTGTAGAGAGTG	1140
Dp	1081	CACGAGCTGGGCACTGTTCACAATGACCCCAATGACAAATGTAAAGTCTGTAGAGAGTG	1140


```

; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: us/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7473089CB1
US-10-311-035-32

```

```

Query Match      77.0%; Score 2196; DB 9; Length 2930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

```

QY 1 ATGCTTGTGCTGGGATCCTTAACCTGCTTTCGCGGGGAAACCGCTGAGAGCTCTAG 60
DB 75 ATGCTTGTGCTGGGATCCTTAACCTGCTTTCGCGGGGAAACCGCTGAGAGCTCTAG 134
QY 61 CCAGAGCGGAGGTAGTCTGTTCCCATCGACTGGACCCCGACATTAACGCGCGCTAC 120
DB 135 CCAGAGCGGAGGTAGTCTGTTCCCATCGACTGGACCCCGACATTAACGCGCGCTAC 194
QY 121 TACTGCGGGGTCGCGAGAGCTCCGGGATCCAGGAGCTATTTTCAGATCAGAGATT 180
DB 195 TACTGCGGGGTCGCGAGAGCTCCGGGATCCAGGAGCTATTTTCAGATCAGAGATT 254
QY 181 CAGAGAGCTTTTACCTACACCTGAGCGGAGATGCTCACTTCTTGAGCTCCCGCTTCTCC 240
DB 255 CAGAGAGCTTTTACCTACACCTGAGCGGAGATGCTCACTTCTTGAGCTCCCGCTTCTCC 314
QY 241 ACTGAGCACTGTGGGCTCCCTCCCTCAAGGGGCTACCGGGGGCTCTTCAGACCTGAGAGC 300
DB 315 ACTGAGCACTGTGGGCTCCCTCCCTCAAGGGGCTACCGGGGGCTCTTCAGACCTGAGAGC 374
QY 301 TGGTTTATTTGGGGAGCTGAAACGCGGAGCCGAGACTGCTTGGCTGTCTGTGAGCTGTGC 360
DB 375 TGGTTTATTTGGGGAGCTGAAACGCGGAGCCGAGACTGCTTGGCTGTCTGTGAGCTGTGC 434
QY 361 GGGGGGCTCCGCGGAGACTTTGGCTTACCGAGGCGCGAGATATGTATTAGCCGCTGCC 420
DB 435 GGGGGGCTCCGCGGAGACTTTGGCTTACCGAGGCGCGAGATATGTATTAGCCGCTGCC 494
QY 421 AATGCTAGCGCGCGCGGCGGAGCGCAACAGCGAGGGCGCACACCTTCTCCAGCGCGG 480
DB 495 AATGCTAGCGCGCGCGGCGGAGCGCAACAGCGAGGGCGCACACCTTCTCCAGCGCGG 554
QY 481 GGTGTTCCGGGGGGGCTTCGCGAGACCCCACTCTCGCTGGGGGGTGGCTCGGGGCTGG 540
DB 555 GGTGTTCCGGGGGGGCTTCGCGAGACCCCACTCTCGCTGGGGGGTGGCTCGGGGCTGG 614
QY 541 AACCCCGCATCTCAAGGCGCTTGAACCTTACAAAGCCGCGGCGGGGCTTTCGGGGAG 600
DB 615 AACCCCGCATCTCAAGGCGCTTGAACCTTACAAAGCCGCGGCGGGGCTTTCGGGGAG 674
QY 601 AGTCGTAGCGCGCGGAGCTGTGGGCGCGCAAGCGTTTCGTCTATCCCGGTAGCGTG 660
DB 675 AGTCGTAGCGCGCGGAGCTGTGGGCGCGCAAGCGTTTCGTCTATCCCGGTAGCGTG 734
QY 661 GAGAGCGGTGTGTGCGGAGAGTCAATGTCAAGTTCCAGCGGGGAGGAGCTGGAACAT 720
DB 735 GAGAGCGGTGTGTGCGGAGAGTCAATGTCAAGTTCCAGCGGGGAGGAGCTGGAACAT 794
QY 721 TATGCTGACGCTCTGCGCAACGCGGCGGAGCTCTACCGGCATCCAGCATCTCTAAC 780
DB 795 TATGCTGACGCTCTGCGCAACGCGGCGGAGCTCTACCGGCATCCAGCATCTCTAAC 854
QY 781 CCCATCAACATGTTGTGTGTAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCGCAAG 840

```

```

DB 855 CCCATCAACATGTTGTGTGTAAGGTGCTGCTTCTTAGAGATCGTACTCCGGGCGCAAG 914
QY 841 GTACCCGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTGCGAGAGAAGCTGAAAC 900
DB 915 GTACCCGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTGCGAGAGAAGCTGAAAC 974
QY 901 AAAGTGAAGTGAACAGCACCAGGAGTGGAGACACTGCACTCTTCTTACCAGGAGAGAC 960
DB 975 AAAGTGAAGTGAACAGCACCAGGAGTGGAGACACTGCACTCTTCTTACCAGGAGAGAC 1034
QY 961 CTGTGTGAGGACCAACCACTGTGACACCTGAGGAGCTGATGTGTGAGTACCATGTGAGAC 1020
DB 1035 CTGTGTGAGGACCAACCACTGTGACACCTGAGGAGCTGATGTGTGAGTACCATGTGAGAC 1094
QY 1021 CCCAAGACAAGTGTCTGTCTATTTAGAGAGCATGGGCTTCCATCATGACCTTCCACCTGCC 1080
DB 1095 CCCAAGACAAGTGTCTGTCTATTTAGAGAGCATGGGCTTCCATCATGACCTTCCACCTGCC 1154
QY 1081 CACGAGCTGGGGCGACGTGTCAACATGCCCATGACAAATGTGAAGTCTGTGAGAGAGTG 1140
DB 1155 CACGAGCTGGGGCGACGTGTCAACATGCCCATGACAAATGTGAAGTCTGTGAGAGAGTG 1214
QY 1141 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGAACCTCATCCAGATGAGACCGTCC 1200
DB 1215 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGAACCTCATCCAGATGAGACCGTCC 1274
QY 1201 AACCCCTGTGACGCTGTGATGCTGCATCATCAACCGACTCTCTGGAACGCGGGGACGCT 1260
DB 1275 AACCCCTGTGACGCTGTGATGCTGCATCATCAACCGACTCTCTGGAACGCGGGGACGCT 1334
QY 1261 GACTGCTCTGCGGACCAACCGCAAGGCCATCTCCCTCCCGAGAGATCTGCGGGGCGC 1320
DB 1335 GACTGCTCTGCGGACCAACCGCAAGGCCATCTCCCTCCCGAGAGATCTGCGGGGCGC 1394
QY 1321 AGCTACACCTGTGAGGACCAAGTGTGAGAGCTTTTGGCGGGGCTCCAAAGCCCTGTCTCT 1380
DB 1395 AGCTACACCTGTGAGGACCAAGTGTGAGAGCTTTTGGCGGGGCTCCAAAGCCCTGTCTCT 1454
QY 1381 TACATGCACTGTGACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGTGTGC 1440
DB 1455 TACATGCACTGTGACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGTGTGC 1514
QY 1441 CAGACCGGCACTTCCCTGCGGAGGATGGGACGACGCTGTGGCGGAGCAAGCTTGTCTC 1500
DB 1515 CAGACCGGCACTTCCCTGCGGAGGATGGGACGACGCTGTGGCGGAGCAAGCTTGTCTC 1574
QY 1501 AAAGGGGCTCTGCGAGAGACACACACCTCAACAGACAGGGGTGATGGTTCTGTGGGCG 1560
DB 1575 AAAGGGGCTCTGCGAGAGACACACACCTCAACAGACAGGGGTGATGGTTCTGTGGGCG 1634
QY 1561 AAATGGGATCTCTTATGCGCCCTGCTCGCGACATGTGTGGGGGCTGTGACGTGGCCAG 1620
DB 1635 AAATGGGATCTCTTATGCGCCCTGCTCGCGACATGTGTGGGGGCTGTGACGTGGCCAG 1694
QY 1621 AGGCAAGTGAACCAACCCACCCCTGCGCAAGCGGGGCAAGTACGCGAGGAGTGAAGGTG 1680
DB 1695 AGGCAAGTGAACCAACCCACCCCTGCGCAAGCGGGGCAAGTACGCGAGGAGTGAAGGTG 1754
QY 1681 AAATACGATCTCTCAATCTGAGACCTTGCCCAAGCTCAAGCTCCGGAAGAGCTTCCGG 1740
DB 1755 AAATACGATCTCTCAATCTGAGACCTTGCCCAAGCTCAAGCTCCGGAAGAGCTTCCGG 1814
QY 1741 GAGGAGAGTGTGAGGCTTTCACAGGCTTAACACGACGCAACCGGCTCACTCTCGCC 1800
DB 1815 GAGGAGAGTGTGAGGCTTTCACAGGCTTAACACGACGCAACCGGCTCACTCTCGCC 1874
QY 1801 GTGCGATGAGTGTGCGCAAGTACTCGGCGGTGTCTCCCGGAGACAAGTGAAGCTCATCTGC 1860
DB 1875 GTGCGATGAGTGTGCGCAAGTACTCGGCGGTGTCTCCCGGAGACAAGTGAAGCTCATCTGC 1934
QY 1861 CGAGCAATGAGCACTGCTACTTCTATGTGCTGGACCCAA--GGTGTGACGCGGACAG 1917

```

Db 1935 CGAGCCAAATGGCACTGGCTACTTCTATGTGCTGGACCCCAAGTGGTGTGAGCGGACG 1994

QY 1918 CTGTGCTCTCTGACTCTCACCCTCCGCTGTGTGTCCAAAGGCAAGTCATCAAGSCTGGCTGT 1977

Db 1995 GTGTGCTCTCTGACTCTCACCCTCCGCTGTGTGTCCAAAGGCAAGTCATCAAGSCTGGCTGT 2054

QY 1978 GATGGGAACCTGGGCTCCAGAAAGATTCAGCAAGTGTGGGTGTGTGGGGAGAGCAAT 2037

Db 2055 GATGGGAACCTGGGCTCCAGAAAGATTCAGCAAGTGTGGGTGTGTGGGGAGAGCAAT 2114

QY 2038 AAGAGCTGCAAGAGTGTGACTGACTCTTCAACCAAGCCATGATGAGTCAATTTCTGTG 2097

Db 2115 AAGAGCTGCAAGAGTGTGACTGACTCTTCAACCAAGCCATGATGAGTCAATTTCTGTG 2174

QY 2098 GTGGCCATCCCGGAGGCGCTCAAGCATGACATCCGCGAGCGGTTTCAAGGGCTG 2157

Db 2175 GTGGCCATCCCGGAGGCGCTCAAGCATGACATCCGCGAGCGGTTTCAAGGGCTG 2234

QY 2158 ATCGGGGATGCAACTACTGCTGCTTGAAGACAGCCCAAGCAAGTACTGCTCAACGGG 2217

Db 2235 ATCGGGGATGCAACTACTGCTGCTTGAAGACAGCCCAAGCAAGTACTGCTCAACGGG 2294

QY 2218 CATTTCTGTGTGTGGGGGTGTGAGCGGAGCTGTGTGTAAGGCACTGTCTGCGGTAC 2277

Db 2295 CATTTCTGTGTGTGGGGGTGTGAGCGGAGCTGTGTGTAAGGCACTGTCTGCGGTAC 2354

QY 2278 AGCGGACAGGGGACAGGCGGTGTGAGAGCGCTTCCGCGCCCATCTTGAAGCGGCTG 2337

Db 2355 AGCGGACAGGGGACAGGCGGTGTGAGAGCGCTTCCGCGCCCATCTTGAAGCGGCTG 2414

QY 2338 ACCGTGAGAGTCTCTCTCGTGGGAGAGATGACACCGGCGGTGCGCTACTCTTCTAT 2397

Db 2415 ACCGTGAGAGTCTCTCTCGTGGGAGAGATGACACCGGCGGTGCGCTACTCTTCTAT 2474

QY 2398 CTGGCCAAAGAGCTCGGAGAGCAAGTCTCTATCCC 2436

Db 2475 CTGGCCAAAGAGCTCGGAGAGCAAGTCTCTATCCC 2513

RESULT 8

US-60-453-135-77889

Sequence 77889, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOUBOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CI001456

CURRENT APPLICATION NUMBER: US/60/453,135

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 77889

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-60-453-135-77889

Query Match

Best Local Similarity 5.3%; Score 150; DB 11; Length 201;

Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 138 GGACTCCGGGATGAGGACTCATTTTTCAGATCAGACATTTTCAGAGGAGCTTTTACT 197

Db 1 GGACTCCGGGATGAGGACTCATTTTTCAGATCAGACATTTTCAGAGGAGCTTTTACT 60

QY 198 ACACCTGACGCGGATGCTAGTCTTGGCTCCGCTTCTCCACTGAGATCTGGGCGT 257

Db 61 ACACCTGACGCGGATGCTAGTCTTGGCTCCGCTTCTCCACTGAGATCTGGGCGT 120

QY 258 CCCCCCTCAGGGGCTCAGCGGGGCTCTTTCAGACCTGGAGAGCGTCTTATTTGGGGA 317

Db 121 CCCCCCTCAGGGGCTCAGCGGGGCTCTTTCAGACCTGGAGAGCGTCTTATTTGGGGA 180

QY 318 CGTGAAGCGGAGCGGAGCT 338

Db 181 CGTGAAGCGGAGCGGAGCTC 201

RESULT 9

US-60-453-135-77890

Sequence 77890, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOUBOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CI001456

CURRENT APPLICATION NUMBER: US/60/453,135

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 77890

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-60-453-135-77890

Query Match

Best Local Similarity 5.3%; Score 150; DB 11; Length 201;

Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 GACTCTTACCAAGCCCATGCTGCTCAATTTGCTGGCCATCCCGGAGCGCT 2119

Db 1 GACTCTTACCAAGCCCATGCTGCTCAATTTGCTGGCCATCCCGGAGCGCT 60

QY 2120 CAAGCATGACATCCGCGGAGCGGTTACAAGGCGCTGATGCGGATGACAATCTG 2179

Db 61 CAAGCATGACATCCGCGGAGCGGTTACAAGGCGCTGATGCGGATGACAATCTG 120

QY 2180 CTCTGAGAACAGCCAGGAGCAAGTCTGCTAACGGGCTTCTGTGTGCGGCTG 2239

Db 121 CTCTGAGAACAGCCAGGAGCAAGTCTGCTAACGGGCTTCTGTGTGCGGCTG 180

QY 2240 AGCGGAGCTGTGGTGTGAAG 2260

Db 181 AGCGGAGCTGTGGTGTGAAG 201

RESULT 10

US-60-453-135-77891

Sequence 77891, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOUBOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CI001456

CURRENT APPLICATION NUMBER: US/60/453,135

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 77891

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-60-453-135-77891

Query Match

Best Local Similarity 5.3%; Score 150; DB 11; Length 201;

Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1598 GTGGGGGCTGACGTGCGCAGAGGACAGTGCACCAACCCCTGCAACGSGGGA 1657

Db 1 GTGGGGGCTGACGTGCGCAGAGGACAGTGCACCAACCCCTGCAACGSGGGA 60


```
QY 1658 AGTACTGCGAGGAGTGTAGGGTGAATACCATCTGCAATCTGGAGCCCTGGCCAGCT 1717
      |||
Db 61 AGTACTGCGAGGAGTGTAGGGTGAATACCATCTGCAATCTGGAGCCCTGGCCAGCT 120
QY 1718 CACCTTCGCGAAAGAGCTTCCGGAGAGACAGTGTAGGCTTTCAACGGCTCAACACACA 1777
      |||
Db 121 CACCTTCGCGAAAGAGCTTCCGGAGAGACAGTGTAGGCTTTCAACGGCTCAACACACA 180
QY 1778 GCACCAACCGGCTCACTCTG 1798
      |||
Db 181 GCACCAACCGGCTCACTCTG 201

RESULT 11
US-60-453-135-77893
; Sequence 77893, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77893
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-77893

Query Match
Best Local Similarity 9.5%; Score 150; DB 11; Length 201;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1768 TACACACAGCAGCACCAACCGGCTCACTCTGCGGTGGCATGGGTGCCCAAGTACTCCGCG 1827
      |||
Db 1 TACACACAGCAGCACCAACCGGCTCACTCTGCGGTGGCATGGGTGCCCAAGTACTCCGCG 60
QY 1828 GTGTCTCCCGGGAGACAAGTGAAGTCATCTCCGAGCAATGAGACACTGTCTTAT 1887
      |||
Db 61 GTGTCTCCCGGGAGACAAGTGAAGTCATCTCCGAGCAATGAGACACTGTCTTAT 120
QY 1888 GTGTGCGACCCAGAGTGTGTGAGCGACAGCTGTCTCTCTGAGTCCACCTCCGCTGT 1947
      |||
Db 121 GTGTGCGACCCAGAGTGTGTGAGCGACAGCTGTCTCTCTGAGTCCACCTCCGCTGT 180
QY 1948 GTCCAAAGCAAGTGCATCAAG 1968
      |||
Db 181 GTCCAAAGCAAGTGCATCAAG 201

RESULT 12
US-60-453-135-77894
; Sequence 77894, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77894
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-77894

Query Match 5.3%; Score 150; DB 11; Length 201;
```

```
Best Local Similarity 99.5%; Pred. No. 5.1e-65;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2186 AGAACGCGCAAGGCAAGTACTGCTCAACAGGGGATTTGCTGTGCGCGGTGAGCGG 2245
      |||
Db 1 AGAACGCGCAAGGCAAGTACTGCTCAACAGGGGATTTGCTGTGCGCGGTGAGCGG 60
QY 2246 ACCTGTGTGAAGGGCAGTGTCTGCGGTACACGGGACAGCGGACAGCGGTGAGAGCC 2305
      |||
Db 61 ACCTGTGTGAAGGGCAGTGTCTGCGGTACACGGGACAGCGGACAGCGGTGAGAGCC 120
QY 2306 TGCAGGCTTCCCGGCCATCTCTGAGCGGTGACCGTGGAGAGTCTCCGTGGGAGAGA 2365
      |||
Db 121 TGCAGGCTTCCCGGCCATCTCTGAGCGGTGACCGTGGAGAGTCTCCGTGGGAGAGA 180
QY 2366 TGACACCGCCCGGGTCCGCT 2386
      |||
Db 181 TGACACCGCCCGGGTCCGCT 201

RESULT 13
US-60-453-050-77889
; Sequence 77889, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77889
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-77889

Query Match
Best Local Similarity 99.5%; Score 150; DB 11; Length 201;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 138 GGACTCCGGGGATGAGGAGTCAATTTTCAGATCAGCATTCAGAGGACTTTTACCT 197
      |||
Db 1 GGACTCCGGGGATGAGGAGTCAATTTTCAGATCAGCATTCAGAGGACTTTTACCT 60
QY 198 ACACCTGACGCGGATGCTCAGTCTTGCTCCCGCTTCTCCACTGAGCATCTGGCGCT 257
      |||
Db 61 ACACCTGACGCGGATGCTCAGTCTTGCTCCCGCTTCTCCACTGAGCATCTGGCGCT 120
QY 258 CCCCCCTCAGGGGCTCACCGGGGGCTTTTCAGACCTGCGACGCTCTTCTATTCTGAGGA 317
      |||
Db 121 CCCCCCTCAGGGGCTCACCGGGGGCTTTTCAGACCTGCGACGCTCTTCTATTCTGAGGA 180
QY 318 CGTGAAGCGGAGCGGAGCTC 338
      |||
Db 181 CGTGAAGCGGAGCGGAGCTC 201

RESULT 14
US-60-453-050-77890
; Sequence 77890, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
```

; SEQ ID NO 77890
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-77890

Query Match 5.3%; Score 150; DB 11; Length 201;
Best Local Similarity 99.5%; Pred. No. 5.1e-65;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 GACTCTTACCACAGCCCATGATGCTACAATTTCTGTGGCCATCCCGCAGGCGCT 2119
|||
Db 1 GACTCTTACCACAGCCCATGATGCTACAATTTCTGTGGCCATCCCGCAGGCGCT 60
QY 2120 CAACGATCGACATCCGCGAGCGGTTACAAGGCGTATCGGGGATGACACTACTG 2179
|||
Db 61 CAACGATCGACATCCGCGAGCGGTTACAAGGCGTATGCGGATGACACTACTG 120
QY 2180 CTCTGAAGAACAGCCAGGCAAGTACCTGCTCAACGGGCAATTCGTGTGCGGCGTGG 2239
|||
Db 121 CTCTGAAGAACAGCCAGGCAAGTACCTGCTCAACGGGCAATTCGTGTGCGGCGTGG 180
QY 2240 AGCGGACCTGTGTGTGAAG 2260
|||
Db 181 AGCGGACCTGTGTGTGAAG 201

RESULT 15

US-60-453-050-77891
; Sequence 77891, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CU001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77891
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-77891

Query Match 5.3%; Score 150; DB 11; Length 201;

Best Local Similarity 99.5%; Pred. No. 5.1e-65;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1598 GTGGGGGCTGACCTGGCCAGAGGAGTGCACCAACCCACCCCTGCGCAGGGGCA 1657
|||
Db 1 GTGGGGGCTGACCTGGCCAGAGGAGTGCACCAACCCACCCCTGCGCAGGGGCA 60
QY 1658 AGTACTGGAGGAGTGAAGGTGAATACGATCTGCAATCTGAGAGCCCTGCCAGCT 1717
|||
Db 61 AGTACTGGAGGAGTGAAGGTGAATACGATCTGCAATCTGAGAGCCCTGCCAGCT 120
QY 1718 CAGCCTCGGAAAGGCTTCGCGGAGGAGAGTGTGAGGCTTCAACGGCTACACACCA 1777
|||
Db 121 CAGCCTCGGAAAGGCTTCGCGGAGGAGAGTGTGAGGCTTCAACGGCTACACACCA 180
QY 1778 GCACCAACCGGCTCCTCTCG 1798
|||
Db 181 GCACCAACCGGCTCCTCTCG 201

Search completed: May 16, 2003, 09:12:03
Job time : 12847 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 19:52:06 ; Search time 5576 Seconds
(without alignments)
12664.332 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853
Sequence: 1 atgctctgctgggacatcct.....gcgtctgagggctgctga 2853

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents, NA, Main:*

```
1: /cgcn2_6/ptodata/1/pna/US0955_COMB.seq:*
2: /cgcn2_6/ptodata/1/pna/US0956_COMB.seq:*
3: /cgcn2_6/ptodata/1/pna/US0957_COMB.seq:*
4: /cgcn2_6/ptodata/1/pna/US0958_COMB.seq:*
5: /cgcn2_6/ptodata/1/pna/US0959_COMB.seq:*
6: /cgcn2_6/ptodata/1/pna/US0960_COMB.seq:*
7: /cgcn2_6/ptodata/1/pna/US0961_COMB.seq:*
8: /cgcn2_6/ptodata/1/pna/US0962_COMB.seq:*
9: /cgcn2_6/ptodata/1/pna/US0963_COMB.seq:*
10: /cgcn2_6/ptodata/1/pna/US0964_COMB.seq:*
11: /cgcn2_6/ptodata/1/pna/US0965_COMB.seq:*
12: /cgcn2_6/ptodata/1/pna/US0966_COMB.seq:*
13: /cgcn2_6/ptodata/1/pna/US0967_COMB.seq:*
14: /cgcn2_6/ptodata/1/pna/US0968_COMB.seq:*
15: /cgcn2_6/ptodata/1/pna/US0969_COMB.seq:*
16: /cgcn2_6/ptodata/1/pna/US0970_COMB.seq:*
17: /cgcn2_6/ptodata/1/pna/US0971_COMB.seq:*
18: /cgcn2_6/ptodata/1/pna/US0972_COMB.seq:*
19: /cgcn2_6/ptodata/1/pna/US0973_COMB.seq:*
20: /cgcn2_6/ptodata/1/pna/US0974_COMB.seq:*
21: /cgcn2_6/ptodata/1/pna/US0975_COMB.seq:*
22: /cgcn2_6/ptodata/1/pna/US0976_COMB.seq:*
23: /cgcn2_6/ptodata/1/pna/US0977_COMB.seq:*
24: /cgcn2_6/ptodata/1/pna/US0978_COMB.seq:*
25: /cgcn2_6/ptodata/1/pna/US0979_COMB.seq:*
26: /cgcn2_6/ptodata/1/pna/US0980_COMB.seq:*
27: /cgcn2_6/ptodata/1/pna/US0981_COMB.seq:*
28: /cgcn2_6/ptodata/1/pna/US0982_COMB.seq:*
29: /cgcn2_6/ptodata/1/pna/US0983_COMB.seq:*
30: /cgcn2_6/ptodata/1/pna/US0984_COMB.seq:*
31: /cgcn2_6/ptodata/1/pna/US0985_COMB.seq:*
32: /cgcn2_6/ptodata/1/pna/US0986_COMB.seq:*
33: /cgcn2_6/ptodata/1/pna/US0987_COMB.seq:*
34: /cgcn2_6/ptodata/1/pna/US0988_COMB.seq:*
35: /cgcn2_6/ptodata/1/pna/US0989_COMB.seq:*
36: /cgcn2_6/ptodata/1/pna/US0990_COMB.seq:*
37: /cgcn2_6/ptodata/1/pna/US0991_COMB.seq:*
38: /cgcn2_6/ptodata/1/pna/US0992_COMB.seq:*
39: /cgcn2_6/ptodata/1/pna/US0993_COMB.seq:*
40: /cgcn2_6/ptodata/1/pna/US0994_COMB.seq:*
41: /cgcn2_6/ptodata/1/pna/US0995_COMB.seq:*
42: /cgcn2_6/ptodata/1/pna/US0996_COMB.seq:*
43: /cgcn2_6/ptodata/1/pna/US0997_COMB.seq:*
```

SUMMARIES			
Result No.	Score	Query Length DB	ID
1	2853	100.0	US-09-965-631-3
2	2853	100.0	US-09-965-631-7
3	2848.2	99.8	US-09-741-151-1
4	2846.6	99.8	US-10-009-332-2
5	2817.6	98.8	US-60-216-821-9*
6	2619.2	91.8	US-10-093-463-27
7	2299.4	80.6	US-60-360-207-12354
8	2298	80.5	US-10-163-316-3
9	2298	80.5	US-60-297-863-3
10	2298	80.5	US-10-163-316-1
11	2298	80.5	US-60-297-863-1
12	2285.4	78.0	US-60-212-656-762
13	2285.4	80.1	US-60-242-679-1814
14	1997.8	66.9	US-60-230-435-2152
15	1671.4	58.6	US-09-575-003-17
16	1285.8	45.1	US-09-609-059-2
17	1285.8	45.1	US-09-975-545-17
18	1285.8	45.1	US-10-071-241-2
19	1091.2	38.2	US-09-965-631-5
20	960.4	33.7	US-60-206-028-17
21			
44:	/cgcn2_6/ptodata/1/pna/US6000_COMB.seq:*		
45:	/cgcn2_6/ptodata/1/pna/US6001_COMB.seq:*		
46:	/cgcn2_6/ptodata/1/pna/US6002_COMB.seq:*		
47:	/cgcn2_6/ptodata/1/pna/US6003_COMB.seq:*		
48:	/cgcn2_6/ptodata/1/pna/US6004_COMB.seq:*		
49:	/cgcn2_6/ptodata/1/pna/US6005_COMB.seq:*		
50:	/cgcn2_6/ptodata/1/pna/US6006_COMB.seq:*		
51:	/cgcn2_6/ptodata/1/pna/US6007_COMB.seq:*		
52:	/cgcn2_6/ptodata/1/pna/US6008_COMB.seq:*		
53:	/cgcn2_6/ptodata/1/pna/US6009_COMB.seq:*		
54:	/cgcn2_6/ptodata/1/pna/US6010_COMB.seq:*		
55:	/cgcn2_6/ptodata/1/pna/US6011_COMB.seq:*		
56:	/cgcn2_6/ptodata/1/pna/US6012_COMB.seq:*		
57:	/cgcn2_6/ptodata/1/pna/US6013_COMB.seq:*		
58:	/cgcn2_6/ptodata/1/pna/US6014_COMB.seq:*		
59:	/cgcn2_6/ptodata/1/pna/US6015_COMB.seq:*		
60:	/cgcn2_6/ptodata/1/pna/US6016_COMB.seq:*		
61:	/cgcn2_6/ptodata/1/pna/US6017_COMB.seq:*		
62:	/cgcn2_6/ptodata/1/pna/US6018_COMB.seq:*		
63:	/cgcn2_6/ptodata/1/pna/US6019_COMB.seq:*		
64:	/cgcn2_6/ptodata/1/pna/US6020_COMB.seq:*		
65:	/cgcn2_6/ptodata/1/pna/US6021_COMB.seq:*		
66:	/cgcn2_6/ptodata/1/pna/US6022_COMB.seq:*		
67:	/cgcn2_6/ptodata/1/pna/US6023_COMB.seq:*		
68:	/cgcn2_6/ptodata/1/pna/US6024_COMB.seq:*		
69:	/cgcn2_6/ptodata/1/pna/US6025_COMB.seq:*		
70:	/cgcn2_6/ptodata/1/pna/US6026_COMB.seq:*		
71:	/cgcn2_6/ptodata/1/pna/US6027_COMB.seq:*		
72:	/cgcn2_6/ptodata/1/pna/US6028_COMB.seq:*		
73:	/cgcn2_6/ptodata/1/pna/US6029_COMB.seq:*		
74:	/cgcn2_6/ptodata/1/pna/US6030_COMB.seq:*		
75:	/cgcn2_6/ptodata/1/pna/US6031_COMB.seq:*		
76:	/cgcn2_6/ptodata/1/pna/US6032_COMB.seq:*		
77:	/cgcn2_6/ptodata/1/pna/US6033_COMB.seq:*		
78:	/cgcn2_6/ptodata/1/pna/US6034_COMB.seq:*		
79:	/cgcn2_6/ptodata/1/pna/US6035_COMB.seq:*		
80:	/cgcn2_6/ptodata/1/pna/US6036_COMB.seq:*		
81:	/cgcn2_6/ptodata/1/pna/US6037_COMB.seq:*		
82:	/cgcn2_6/ptodata/1/pna/US6038_COMB.seq:*		
83:	/cgcn2_6/ptodata/1/pna/US6039_COMB.seq:*		
84:	/cgcn2_6/ptodata/1/pna/US6040_COMB.seq:*		
85:	/cgcn2_6/ptodata/1/pna/US6041_COMB.seq:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

QY	301	TGCTTCATATTCGGGGAGCTAAAGCCGAGCCGAGCTGTTCCTGCTGTGAGCGCTGTC	360
Db	301	TGCTTCATATTCGGGGAGCTAAAGCCGAGCCGAGCTGTTCCTGCTGTGAGCGCTGTC	360
QY	361	GGGGGGCTCCGGAGCCTTTGGCTACCGAGGGCCGAGTATGTATTAAGCCGCTGCC	420
Db	361	GGGGGGCTCCCGAGCCTTTGGCTACCGAGGGCCGAGTATGTATTAAGCCGCTGCC	420
QY	421	AATGCTACCGCGCGGGGGGGCGAGCGCAACAGCCAGGGGCGACACCTTCTTCACAGCCCG	480
Db	421	AATGCTACCGCGCGGGGGGGCGAGCGCAACAGCCAGGGGCGACACCTTCTTCACAGCCCG	480
QY	481	GGTGTTCGGGGGGGCTTCGCGAGACCCACCTCTCGTCCGGGGTGGCTCGGCTGG	540
Db	481	GGTGTTCGGGGGGGCTTCGCGAGACCCACCTCTCGTCCGGGGTGGCTCGGCTGG	540
QY	541	AACCCCGCATCTCAAGGGGCGCTGAGACCCCTTAACAAGCGGGGGGGGGCTTGCGGGAG	600
Db	541	AACCCCGCATCTCAAGGGGCGCTGAGACCCCTTAACAAGCGGGGGGGGGCTTGCGGGAG	600
QY	601	AGTCGTAGCGCGCAGGCTGTGGGGCGCCCAAGGTTTCGTGTATCCCGCGTACGTG	660
Db	601	AGTCGTAGCGCGCAGGCTGTGGGGCGCCCAAGGTTTCGTGTATCCCGCGTACGTG	660

07/15

Db	721	TATCTGCTGACGCTCTGTGGCAACGGCGGGGACCTTACGCCCATCCGACATCTTCAAC	781
Qy	781	CCATCAACATGTTGTGTGTCAAAGTGTGCTTCTTAGATATCGATCTCCGGGCCAAG	840
Db	781	CCATCAACATGTTGTGTGTCAAAGTGTGCTTCTTAGATATCGATCTCCGGGCCAAG	840
Qy	841	GTACACGGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTTGGCAGAGAAAGCTGAC	900
Db	841	GTACACGGCAATGCGGCCCTGACGCTGCGCAACTTCTGTGCTTGGCAGAGAAAGCTGAC	900
Qy	901	AAAGTAGAGTAACGACACCCCGAGTACAGTGGGACAATGGCAATCTCTTACGAGCGAGAC	960
Db	901	AAAGTAGAGTAACGACACCCCGAGTACAGTGGGACAATGGCAATCTCTTACGAGCGAGAC	960
Qy	961	CTGTGTGAGGACCAACACCTGTGACACCCCTGGGCATGGCTGATGTGGGTACATGTGTAC	1020
Db	961	CTGTGTGAGGACCAACACCTGTGACACCCCTGGGCATGGCTGATGTGGGTACATGTGTAC	1020

1081 CACCACTCCGCCGACCGCTGTCACATGTCGCCCATGACACATGTGAAAGTCTGTGAGTGGGTG 114

Db	1081	CACGAGCTGGGCCACTGTTTCAACATGCCCCATGACATGTGAAAGTGTGTGAGAGGTG	11
Qy	1141	TTTGGGAACTCCGAGCCAAACCATGATGTCCCGACCCATATCAGATGAGCCGTGCC	120
Db	1141	TTTGGGAAAGTCCGAGCCAAACCATGATGTGTCCCGACCCATATCAGATGAGCCGTGCC	120
Qy	1201	AACCCCTGTGACAGCTGCAGTGTGCATCATACCGACTTCTTGAGACGCGGGTACGGT	120
Db	1201	AACCCCTGTGACAGCTGCAGTGTGCATCATACCGACTTCTTGAGACGCGGGTACCGT	120
Qy	1261	GACTGCTCTGTGACCAACCCAGCAAGCCCATCTCCGTCCCGAGAGATCTCCGGGGGCC	13
Db	1261	GACTGCTCTGTGACCAACCCAGCAAGCCCATCTCCGTCCCGAGAGATCTCCGGGGGCC	13
Qy	1321	AGCTACACCTGTGAGCCAGCATGTGAGTGTGTTTGGGTGGGGTCCAAAGCCCTGTCTT	13
Db	1321	AGCTACACCTGTGAGCCAGCATGTGAGTGTGTTTGGGTGGGGTCCAAAGCCCTGTCTT	13

Db 697 TGCCTCTATTCTGGGAGCTGAGCCGAGCCGAGCTGCTCCGCTGCTGAGCCTGTGC 756
QY 361 GGGGGGCTCCGGGAGACCTTTGGCTACCGAGGCGCCGAGTATGTCATTATATCCGCTGCCC 420
Db 757 GGGGGGCTCCGGGAGACCTTTGGCTACCGAGGCGCCGAGTATGTCATTATATCCGCTGCCC 816
QY 421 AATGCTAGCGCGCGCGCGCGAGCGCAAGCCAGGCGCGCAACGCTTTCACGCGCGG 480
Db 817 AATGCTAGCGCGCGCGCGCGAGCGCAAGCCAGGCGCGCAACGCTTTCACGCGCGG 876
QY 481 GGTGTTCCGGGCGCGCTTCCGAGAACCCCACTCTCGCTCGGGGTGGCTCGGGCTGG 540
Db 877 GGTGTTCCGGGCGCGCTTCCGAGAACCCCACTCTCGCTCGGGGTGGCTCGGGCTGG 936
QY 541 AACCCCGCATCTACGCGGCGCTGAGACCTTACAAAGCCGCGCGCGGCTTGGGGAG 600
Db 937 AACCCCGCATCTACGCGGCGCTGAGACCTTACAAAGCCGCGCGGCTTGGGGAG 996
QY 601 AGTCGTAGCGCGCGAGGCTGGGGCGCGCAAGCGTTTGTGTATCCCGGCTAGCTG 660
Db 997 AGTCGTAGCGCGCGAGGCTGGGGCGCGCAAGCGTTTGTGTATCCCGGCTAGCTG 1056
QY 661 GAGACGCTGTGTGCGCGAGCATGTCATGTCATGTCAGCGCGGACCTGGACAT 720
Db 1057 GAGACGCTGTGTGCGCGAGCATGTCATGTCATGTCAGCGCGGACCTGGACAT 1116
QY 721 TATCTGCTAGCGCTGTGCGCAAGCGGGGCGAGCTTACCGCATCCGATCCTCTAC 780
Db 1117 TATCTGCTAGCGCTGTGCGCAAGCGGGGCGAGCTTACCGCATCCGATCCTCTAC 1176
QY 781 CCCATCAATGCTGTGTGTCAGAGTGTCTTCTAGAGATGTCAGCTTCCGGGCCAG 840
Db 1177 CCCATCAATGCTGTGTGTCAGAGTGTCTTCTAGAGATGTCAGCTTCCGGGCCAG 1236
QY 841 GTACCGCGCAATGCGGCGCTGAGCGTGTGCTGTGCTGTGAGAGAGAGAGTGTGAC 900
Db 1237 GTACCGCGCAATGCGGCGCTGAGCGTGTGCTGTGCTGTGAGAGAGAGTGTGAC 1296
QY 901 AAGTGTAGTACAAAGACCCCGAGTGTGAGGAGTGTGCTGTGCTGTGAGAGAGAG 960
Db 1297 AAGTGTAGTACAAAGACCCCGAGTGTGAGGAGTGTGCTGTGCTGTGAGAGAGAG 1356
QY 961 CTGTGTGAGCCACCACTGTGTGACACCTGTGAGCGATGGCTGTGATGTGGTACATGTGTGAC 1020
Db 1357 CTGTGTGAGCCACCACTGTGTGACACCTGTGAGCGATGGCTGTGATGTGGTACATGTGTGAC 1416
QY 1021 CCCAAGAGAGTGTGTGTGTCATGAGAGAGTGTGCTGTGCTGTGACACCTGTGAC 1080
Db 1417 CCCAAGAGAGTGTGTGTGTCATGAGAGAGTGTGCTGTGCTGTGACACCTGTGAC 1476
QY 1081 CAGAGCTGGGCGACGTTTCAACATGTCCCATGACATGTGAAGTGTGTGAGAGTGTG 1140
Db 1477 CAGAGCTGGGCGACGTTTCAACATGTCCCATGACATGTGAAGTGTGTGAGAGTGTG 1536
QY 1141 TTTTGGAGAGTCCGAGGCAACACATGATGTCCCGACCTCATTCAGATGTGACGCTGGC 1200
Db 1537 TTTTGGAGAGTCCGAGGCAACACATGATGTCCCGACCTCATTCAGATGTGACGCTGGC 1596
QY 1201 AACCCGTGTGAGCTGAGTGTGCTGAGTGTGCTGTGAGTGTGCTGTGAGTGTGAGTGTG 1260
Db 1597 AACCCGTGTGAGCTGAGTGTGCTGAGTGTGCTGTGAGTGTGCTGTGAGTGTGAGTGTG 1656
QY 1261 GACTGCTCTGTGAGCAACCCAGACAGCCATCTCTCTGCGAGATGTGCGGGGCGCC 1320
Db 1657 GACTGCTCTGTGAGCAACCCAGACAGCCATCTCTCTGCGAGATGTGCGGGGCGCC 1716
QY 1321 AGCTACACCTGTGAGCGCAAGTGTGAGTGTGCTTGTGGCTGGGCTTCCAGACCTGTGCT 1380
Db 1717 AGCTACACCTGTGAGCGCAAGTGTGAGTGTGCTTGTGGCTGGGCTTCCAGACCTGTGCT 1776
QY 1381 TACATGAGTACATGACCAAGCTGTGTGACCGGAGAGCCAAAGGAGACATGTGTGCT 1440
Db 1777 TACATGAGTACATGACCAAGCTGTGTGACCGGAGAGCCAAAGGAGACATGTGTGCT 1836

QY 1441 CAGACCCGCACTTCCCTGGGCGGATGTGACACAGCTGTGGGAGGCGCAAGCTGTGCTG 1500
Db 1837 CAGACCCGCACTTCCCTGGGCGGATGTGACACAGCTGTGGGAGGCGCAAGCTGTGCTG 1896
QY 1501 AAGGGGCTCGGTGAGAGACCAACCTTCAACAGCACAGGCTGTGCTGTGGGCT 1560
Db 1897 AAGGGGCTCGGTGAGAGACCAACCTTCAACAGCACAGGCTGTGCTGTGGGCT 1956
QY 1561 AATGAGATCCATGAGCCCTGCTGGCGACATGTGTGGGGGCTGTGAGCTGTGACAG 1620
Db 1957 AATGAGATCCATGAGCCCTGCTGGCGACATGTGTGGGGGCTGTGAGCTGTGACAG 2016
QY 1621 AGCAGTGTGACCAACCCCACTGTGCAACGCGGGGCAAGTACTGTGAGGAGTGTG 1680
Db 2017 AGCAGTGTGACCAACCCCACTGTGCAACGCGGGGCAAGTACTGTGAGGAGTGTG 2076
QY 1681 AATATCGATCTGTGCAATGTGAGCCCTGTGCCCCAGCTGTGAGGCTGTGAGGCT 1740
Db 2077 AATATCGATCTGTGCAATGTGAGCCCTGTGCCCCAGCTGTGAGGCTGTGAGGCT 2136
QY 1741 GAGGAGAGTGTGAGGCTTCAACGCGCTACACAGCACAGCAACCGGCTCACTCTGCGC 1800
Db 2137 GAGGAGAGTGTGAGGCTTCAACGCGCTACACAGCACAGCAACCGGCTCACTCTGCGC 2196
QY 1801 GTGGCATGGTGTGCCCCAGTACTCCGCGCTGTCTCCCGGGAGCAAGTGTGATCTGCTG 1860
Db 2197 GTGGCATGGTGTGCCCCAGTACTCCGCGCTGTCTCCCGGGAGCAAGTGTGATCTGCTG 2256
QY 1861 CGAGCCAAATGGCACTGTGCTTCTATGTGTGTGAGCAACCAAGTGTGTGAGCGAGCTG 1920
Db 2257 CGAGCCAAATGGCACTGTGCTTCTATGTGTGTGAGCAACCAAGTGTGTGAGCGAGCTG 2216
QY 1921 TGTCTCTGCTGCTGCTGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 2317 TGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2376
QY 1981 GGGAACTGGGCTTCCAGAGAGAGTGTGAGCAAGTGTGGGGTGTGGGGAGAGCAATAAG 2040
Db 2377 GGGAACTGGGCTTCCAGAGAGAGTGTGAGCAAGTGTGGGGTGTGGGGAGAGCAATAAG 2436
QY 2041 AGCTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2100
Db 2437 AGCTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2496
QY 2101 GCCATCCCGGAGGCGCTTCAAGCATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2160
Db 2497 GCCATCCCGGAGGCGCTTCAAGCATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2556
QY 2161 GGGGATGTGACACTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2220
Db 2557 GGGGATGTGACACTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2616
QY 2221 TTTGCTGTGTGCGGCGGTGTGAGCGGAGCTGTGTGTGAGGCGTGTGTGTGTGTGTGTG 2280
Db 2617 TTTGCTGTGTGCGGCGGTGTGAGCGGAGCTGTGTGTGAGGCGTGTGTGTGTGTGTGTG 2676
QY 2281 GGCAGCGGCAAGCGGTGTGAGAGCTGTGAGGCTTCCCGGCTTCCAGTGTGAGGCTGAGC 2340
Db 2677 GGCAGCGGCAAGCGGTGTGAGAGCTGTGAGGCTTCCCGGCTTCCAGTGTGAGGCTGAGC 2736
QY 2341 GTGAGAGTGTGCTGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 2400
Db 2737 GTGAGAGTGTGCTGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 2796
QY 2401 CCCAAAGAGCTGTGAGGAGACAGTGTGTCTATCCCAAGAGACCCCGGGGAGCCCTGTGTC 2460
Db 2797 CCCAAAGAGCTGTGAGGAGACAGTGTGTCTATCCCAAGAGACCCCGGGGAGCCCTGTGTC 2856
QY 2461 TTGACACAGAGTGTCTAGCTTCTTCAACAGGAGTGTGAGAGCGGAGACAGAGGCTT 2520
Db 2857 TTGACACAGAGTGTCTAGCTTCTTCAACAGGAGTGTGAGAGCGGAGACAGAGGCTT 2916

```

OY 2521 GCACGCTGGTGGCTGGAGCTGGGGCCCTGCTCCCGAGCTGCGGCACTGGCTTCAG 2580
    |||||||
DB 2917 GCACGCTGGTGGCTGGAGCTGGGGCCCTGCTCCCGAGCTGCGGCACTGGCTTCAG 2976
    |||||||
OY 2581 AAGCGGGGGGTGGAGCTGGGGGGCTCCGGCCGGGAGCGGCAAGGCTCCCTGCTGATGCA 2640
    |||||||
DB 2977 AAGCGGGGGGTGGAGCTGGGGGGCTCCGGCCGGGAGCGGCAAGGCTCCCTGCTGATGCA 3036
    |||||||
OY 2641 GCCCATGAGCCCGTGGAGACACAAAGCCTGGGGGAGCCCTGCGCCACCTGGAGAGCTAGC 2700
    |||||||
DB 3037 GCCCATGAGCCCGTGGAGACACAAAGCCTGGGGGAGCCCTGCGCCACCTGGAGAGCTAGC 3096
    |||||||
OY 2701 GCTTGTCACTCCCTGCTCCAAAGCTGGGGCCGGGAGATTTCAGAGCGGCTCACTCAAGTGT 2760
    |||||||
DB 3097 GCTTGTCACTCCCTGCTCCAAAGCTGGGGCCGGGAGATTTCAGAGCGGCTCACTCAAGTGT 3156
    |||||||
OY 2761 GTGGGCAAGGAGCGGGCTGCTGGCCCGGGAGCAGTGGCAACTGGACCGCAAGCCCCAG 2820
    |||||||
DB 3157 GTGGGCAAGGAGCGGGCTGCTGGCCCGGGAGCAGTGGCAACTGGACCGCAAGCCCCAG 3216
    |||||||
OY 2821 GAGCTGACTTCTGCTGCTGAGGCGGCTGCTGA 2853
    |||||||
DB 3217 GAGCTGACTTCTGCTGCTGAGGCGGCTGCTGA 3249
    |||||||

```

RESULT 3

```

US-09-741-151-1
; Sequence 1, Application US/09741151
; GENERAL INFORMATION:
; APPLICANT: ZHU, ShiaoPing et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1001005
; CURRENT APPLICATION NUMBER: US/09/741,151
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2867
; TYPE: DNA
; ORGANISM: Human
US-09-741-151-1

```

```

Query Match          99.8%; Score 2848.2; DB 29; Length 2867;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2850; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 ATGCTTGTGCTGGCATCTTAACCTGGCTTCCCGGGGCAACCGCTGAGAGCTGAG 60
    |||||||
DB 4 ATGCTTGTGCTGGCATCTTAACCTGGCTTCCCGGGGCAACCGCTGAGAGCTGAG 63
    |||||||
OY 61 CCAGAGCGGAGAGTAGTGTTCCTCCATCCGACTGGAGCCGAGCATTAAAGGCGCGCTTAC 120
    |||||||
DB 64 CCAGAGCGGAGAGTAGTGTTCCTCCATCCGACTGGAGCCGAGCATTAAAGGCGCGCTTAC 123
    |||||||
OY 121 TACTGGCGGGGTCCCGAGAGACTCCGGGAGATCAGGAGCTCATTTTTCAGATCAAGCACTTT 180
    |||||||
DB 124 TACTGGCGGGGTCCCGAGAGACTCCGGGAGATCAGGAGCACTTTTTCAGATCAAGCACTTT 183
    |||||||
OY 181 CAGGAGAGACTTTTACCTACACTGACGCGGAGATGCTCAGTCTTGAGCTCCGCTTTC 240
    |||||||
DB 184 CAGGAGAGACTTTTACCTACACTGACGCGGAGATGCTCAGTCTTGAGCTCCGCTTTC 243
    |||||||
OY 241 ACTGAGCATCTGGCGGTCCCTCCAGGAGGCTCACCGGGGCTTTTACAGACTTGGAGCGC 300
    |||||||
DB 244 ACTGAGCATCTGGCGGTCCCTCCAGGAGGCTCACCGGGGCTTTTACAGACTTGGAGCGC 303
    |||||||
OY 301 TGCCTTATTTTGGGAGAGTGAAGCGAGCGGAGCTGCTGGCTGCTGAGAGCTTGGC 360
    |||||||
DB 304 TGCCTTATTTTGGGAGAGTGAAGCGAGCGGAGCTGCTGGCTGCTGAGAGCTTGGC 363
    |||||||
OY 361 GGGGGGCTCCGCGAGCCTTTGGCTACCGAGGCGCGGAGTATGTCATTAGCCCGCTGCC 420
    |||||||

```

```

DB 364 GGGGGGCTCCCGAGAGCTTTGGCTACCGAGGCGCGGAGTATGATTACCCGCTGCC 423
    |||||||
OY 421 AATGTACAGCGCGCGCGGCGGACAGGCAAGCCAGAGGGGCGACACCTTCCAGCGCCGG 480
    |||||||
DB 424 AATGTACAGCGCGCGCGGCGGACAGGCAAGCCAGAGGGGCGACACCTTCCAGCGCCGG 483
    |||||||
OY 481 GGTGTTCGGGGGCGGCTTCCGGAGACCCCAAGCTCTGCTGGCGGGGTGAGCTTGAGGCTG 540
    |||||||
DB 484 GGTGTTCGGGGGCGGCTTCCGGAGACCCCAAGCTCTGCTGGCGGGGTGAGCTTGAGGCTG 543
    |||||||
OY 541 AACCCCGCATCTTACAGGCGCTTGAGCCCTTTACAAAGCCGCGGCGGGGCTTGGGAGAG 600
    |||||||
DB 544 AACCCCGCATCTTACAGGCGCTTGAGCCCTTTACAAAGCCGCGGCGGGGCTTGGGAGAG 603
    |||||||
OY 601 AGTGTACCGGGCGAGGCTGTGGGCGGCGCAAGGCTTGTGCTATCCCGGCTACGTG 660
    |||||||
DB 604 AGTGTACCGGGCGGAGGCTGTGGGCGGCGCAAGGCTTGTGCTATCCCGGCTACGTG 663
    |||||||
OY 661 GAGAGCTGTGTGTGCGGAGAGTCAATGATCAAGTTCCAGCGGCGGAGCTGGAACAT 720
    |||||||
DB 664 GAGAGCTGTGTGTGCGGAGAGTCAATGATCAAGTTCCAGCGGCGGAGCTGGAACAT 723
    |||||||
OY 721 TATCTGTGAGCTGTGTGCGGAGAGGCTGCTTACCGCATCCAGCATCTCTCAC 780
    |||||||
DB 724 TATCTGTGAGCTGTGTGCGGAGAGGCTGCTTACCGCATCCAGCATCTCTCAC 783
    |||||||
OY 781 CCCATCAACATCTGTTGTGTCAAAGTGTCTCTTGAAGATGCTGATCTCCGCGCCAG 840
    |||||||
DB 784 CCCATCAACATCTGTTGTGTCAAAGTGTCTCTTGAAGATGCTGATCTCCGCGCCAG 843
    |||||||
OY 841 GTACAGCGGCAATGAGGCGGCTGAGCTGCGCAACTTGTGCTGCGGAGAGAAAGCTGAGC 900
    |||||||
DB 844 GTACAGCGGCAATGAGGCGGCTGAGCTGCGCAACTTGTGCTGCGGAGAGAAAGCTGAGC 903
    |||||||
OY 901 AAAGTGTGACAAAGCAACCCCGAGTACTGGAGCACTGCCATCTCTTACACAGGTAGAGC 960
    |||||||
DB 904 AAAGTGTGACAAAGCAACCCCGAGTACTGGAGCACTGCCATCTCTTACACAGGTAGAGC 963
    |||||||
OY 961 CTGTGTGAGGCGACCACTGTGACACCTGGGAGCTGATGTGGGTATCATGTGATGAT 1020
    |||||||
DB 964 CTGTGTGAGGCGACCACTGTGACACCTGGGAGCTGATGTGGGTATCATGTGATGAT 1023
    |||||||
OY 1021 CCCAAGAGAGGCTGCTGTGTCATTGAGAGAGAGTGGCTTCCATCAGCTTACCACTGCC 1080
    |||||||
DB 1024 CCCAAGAGAGGCTGCTGTGTCATTGAGAGAGAGTGGCTTCCATCAGCTTACCACTGCC 1083
    |||||||
OY 1081 CAGGAGCTGGGCGACAGCTTCAACATGCCCATGACAAATGGAAGTCTGTGAGAGGTG 1140
    |||||||
DB 1084 CAGGAGCTGGGCGACAGCTTCAACATGCCCATGACAAATGGAAGTCTGTGAGAGGTG 1143
    |||||||
OY 1141 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGAACCCTCATCCAGATGAGCGTGC 1200
    |||||||
DB 1144 TTTGGGAAGCTCCGAGCCCAACACATGATGTCCCGAACCCTCATCCAGATGAGCGTGC 1203
    |||||||
OY 1201 AACCCCTGTAGCCTGTGAGAGTGTCCATCATCAAGCACTTCCGAGACCGGGGCAAGGT 1260
    |||||||
DB 1204 AACCCCTGTAGCCTGTGAGAGTGTCCATCATCAAGCACTTCCGAGACCGGGGCAAGGT 1263
    |||||||
OY 1261 GACTGCTCTCTGGAGCAACCCAGCAAGGCCATCTCCCTGCGCGAGAGATCTGCGGGCGCT 1320
    |||||||
DB 1264 GACTGCTCTCTGGAGCAACCCAGCAAGGCCATCTCCCTGCGCGAGAGATCTGCGGGCGCT 1323
    |||||||
OY 1321 AGCTACACCTGTAGAGCAGCAGTGGAGCTTTTGGCGGTGGGCTCCAGCCCTGTGCT 1380
    |||||||
DB 1324 AGCTACACCTGTAGAGCAGCAGTGGAGCTTTTGGCGGTGGGCTCCAGCCCTGTGCT 1383
    |||||||
OY 1381 TACATGACAGTACTGACCAACAGCTGTGTGTCACCGGGAAGGCCAAGGAGAGATGTGTGC 1440
    |||||||
DB 1384 TACATGACAGTACTGACCAACAGCTGTGTGTCACCGGGAAGGCCAAGGAGAGATGTGTGC 1443
    |||||||
OY 1441 CAGACCCGCGCACTTCCCTGGGCGGAGTGGACACAGCTGTGGCGAGGGCAAGCTCTGCTC 1500
    |||||||

```

Db 1444 CAGACCCGCTCTCCCTGGGCGCATGGACACAGCTGTGGCGAGGCACTCTGCTC 1503
OY 1501 AAGGGGCTGGGTGAGAGACAACTCAACAAAGACAGAGGTGATGTTCTCTGAGCC 1560
Db 1504 AAGGGGCTGGGTGAGAGACAACTCAACAAAGACAGAGGTGATGTTCTCTGAGCC 1563
OY 1561 AATGGGATCCCTATGAGCCCTGCTGCGGCACATGTGGTGGGGGCGTGCAGTGGCCAGG 1620
Db 1564 AATGGGATCCCTATGAGCCCTGCTGCGGCACATGTGGTGGGGGCGTGCAGTGGCCAGG 1623
OY 1621 AGCAGTGCACCAACCCACCCCTGCAACAGGGGGCAAGTACTGCGAGGAGTGAAGGCTG 1680
Db 1624 AGCAGTGCACCAACCCACCCCTGCAACAGGGGGCAAGTACTGCGAGGAGTGAAGGCTG 1683
OY 1681 AATATCCGATCTCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCGGAAAGAGCTTCGG 1740
Db 1684 AATATCCGATCTCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCGGAAAGAGCTTCGG 1743
OY 1741 GAGAGCAGTGTGAGGCTTTCACAGGCTACACAGCAGACACAGCCGCTCAGCTCTGCGC 1800
Db 1744 GAGAGCAGTGTGAGGCTTTCACAGGCTACACAGCAGACACAGCCGCTCAGCTCTGCGC 1803
OY 1801 GTGGCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGAGCAAGTGAAGCTCATCTGC 1860
Db 1804 GTGGCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGAGCAAGTGAAGCTCATCTGC 1863
OY 1861 CGAGCCATGGGCACTGGCTACTTCTATGTGTGTCGACCCAAAGGTGTGAGCAGGCTG 1923
Db 1864 CGAGCCATGGGCACTGGCTACTTCTATGTGTGTCGACCCAAAGGTGTGAGCAGGCTG 1923
OY 1921 TGGCTCTGACTTCACACTCCGCTGTGTGTCACAGCAAGTGAAGTGTGCTGTGAT 1980
Db 1924 TGGCTCTGACTTCACACTCCGCTGTGTGTCACAGCAAGTGAAGTGTGCTGTGAT 1983
OY 1981 GGGAACTGGGCTCCAAAGAGATTCGACAGTGTGGGGTGTGTGGGGAGCAATAG 2040
Db 1984 GGGAACTGGGCTCCAAAGAGATTCGACAGTGTGGGGTGTGTGGGGAGCAATAG 2043
OY 2041 AGCTGGCAAGAGTGTGACTGTGACTTTCACCAAGCCATGATGATGATTCATTCATTCGAGTG 2100
Db 2044 AGCTGGCAAGAGTGTGACTGTGACTTTCACCAAGCCATGATGATGATTCATTCATTCGAGTG 2103
OY 2101 GGCATCCCGCAGGCGCTCAAGCATCGACATCCGCGAGCGGTATACAAGGGCTGATC 2160
Db 2104 GGCATCCCGCAGGCGCTCAAGCATCGACATCCGCGAGCGGTATACAAGGGCTGATC 2163
OY 2161 GGGGATACAACTACCTGCTGTGAGAAACAGCCAAAGGCAAGTACTGCTCAACGGGGCAT 2220
Db 2164 GGGGATACAACTACCTGCTGTGAGAAACAGCCAAAGTACTGCTCAACGGGGCAT 2223
OY 2221 TTGCTGTGTGCGGGGTGAGAGCGGACCTGTGTGTAAGGAGTGTGCTGCGGTACAGC 2280
Db 2224 TTGCTGTGTGCGGGGTGAGAGCGGACCTGTGTGTAAGGAGTGTGCTGCGGTACAGC 2283
OY 2281 GGCACGGGACAGCGGTGAGAGCTGAGAGCTTCCGGCCCATCTGGAAGCCGCTGACC 2340
Db 2284 GGCACGGGACAGCGGTGAGAGCTGAGAGCTTCCGGCCCATCTGGAAGCCGCTGACC 2343
OY 2341 GTGAGAGTCCCTCCGTGGGGAATGACACCGCCCGGGGCGCTACTCTCTATCTG 2400
Db 2344 GTGAGAGTCCCTCCGTGGGGAATGACACCGCCCGGGGCGCTACTCTCTATCTG 2403
OY 2401 CCCAAAGAGCTCGGAGAGCAAGTCTCATCCCAAGGACCCCGGGAGCCCTCTGTC 2460
Db 2404 CCCAAAGAGCTCGGAGAGCAAGTCTCATCCCAAGGACCCCGGGAGCCCTCTGTC 2463
OY 2461 TTGCAACAACAGCGTCTCAGCTCTCAACAGGTGAGAGCAGCCGGACAGACAGGCCCT 2520
Db 2464 TTGCAACAACAGCGTCTCAGCTCTCAACAGGTGAGAGCAGCCGGACAGACAGGCCCT 2523
OY 2521 GCAGCGTGGGTGGGAGTGGAGTGGGGGCGTCTCCGAGAGCTGCGGAGTGGGCTGAG 2580
Db 2524 GCAGCGTGGGTGGGAGTGGAGTGGGGGCGTCTCCGAGAGCTGCGGAGTGGGCTGAG 2583

OY 2581 AAGGGGCGGTGAGAGTGTGGGCTCCGCGGGGACGACGAGGTCCCTGCTGTGATGA 2640
Db 2584 AAGGGGCGGTGAGAGTGTGGGCTCCGCGGGGACGACGAGGTCCCTGCTGTGATGA 2643
OY 2641 GCCCATGAGCCCGGTGAGACCAAGCCTGGGGGAGCCCTGCCCCACCTGGAGTCAAC 2700
Db 2644 GCCCATGAGCCCGGTGAGACCAAGCCTGGGGGAGCCCTGCCCCACCTGGAGTCAAC 2703
OY 2701 GCTGTGTACCCCTGCTCTCAAGACTGTGGGCGGGGATTCAGAGGGGCTCAGT 2760
Db 2704 GCTGTGTACCCCTGCTCTCAAGACTGTGGGCGGGGATTCAGAGGGGCTCAGT 2763
OY 2761 GTGGGCCACGAGGCGCGCTGTGCCCCGGGAGCAGTGAAGTTCGACCGCAAGCCCA 2820
Db 2764 GTGGGCCACGAGGCGCGCTGTGCCCCGGGAGCAGTGAAGTTCGACCGCAAGCCCA 2823
OY 2821 GAGCTGAGCTTCTGCTGCTGAGAGCCGCTGGA 2853
Db 2824 GAGCTGAGCTTCTGCTGCTGAGAGCCGCTGGA 2856

RESULT 4
US-10-009-332-2
Sequence 2, Application US/10009332
GENERAL INFORMATION:
APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
APPLICANT: Kazusa DNA Research Institute
TITLE OR INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY
FILE REFERENCE: Q67541
CURRENT APPLICATION NUMBER: US/10/009,332
CURRENT FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: JPA Hei 11-321740
PRIOR FILING DATE: 1999-11-11
PRIOR APPLICATION NUMBER: JPA 2000-144020
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 2853
TYPE: DNA
ORGANISM: Homo sapiens
US-10-009-332-2
Query Match 99.8%; Score 2846.6; DB 38; Length 2853;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2849; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 ATGCTTGTGCTGGGATCCTTAACCTGGCTTTCGCGGGGCAACCGCTGAGAGCTTGAG 60
Db 1 ATGCTTGTGCTGGGATCCTTAACCTGGCTTTCGCGGGGCAACCGCTGAGAGCTTGAG 60
OY 61 CCAGAGCGGAGTAGTGTTCCTCCATCCGACTGACCCCGGACATTACGGCCGCCCTAC 120
Db 61 CCAGAGCGGAGTAGTGTTCCTCCATCCGACTGACCCCGGACATTACGGCCGCCCTAC 120
OY 121 TACTGGCGGGGTCCGAGAGACTCCGGGATCAGGAGCATTTTTCAGATACAGCATTT 180
Db 121 TACTGGCGGGGTCCGAGAGACTCCGGGATCAGGAGCATTTTTCAGATACAGCATTT 180
OY 181 CAGGAGACTTTTACTACACTGACGCGGATGCTAGTTCTTGGCTCCGCTTCC 240
Db 181 CAGGAGACTTTTACTACACTGACGCGGATGCTAGTTCTTGGCTCCGCTTCC 240
OY 241 ACTGAGCATGTGGGCTCCCTCCAGAGGCTCACCGGGGCTCTTCAGAGCTGAGAGC 300
Db 241 ACTGAGCATGTGGGCTCCCTCCAGAGGCTCACCGGGGCTCTTCAGAGCTGAGAGC 300
OY 301 TGTCTTATTTCTGGGAGCTGAACCGCGGAGCTGTTGCTGCTGAGCCTGTGC 360
Db 301 TGTCTTATTTCTGGGAGCTGAACCGCGGAGCTGTTGCTGCTGAGCCTGTGC 360
OY 361 GGGGCGCTCCGCGGAGCTTTTGGCTAACGAGGCGCGAGTATGATTAGCCCGCTGCC 420

QY 2581 AAGCGGCGGTGACTGTGCGGGCTCCGCGGCGACGCGTCCCTGCTGTGATGCA 2640
Db 2581 AAGCGGCGGTGACTGTGCGGGCTCCGCGGCGACGCGTCCCTGCTGTGATGCA 2640
QY 2641 GCCATCGGCGGTGAGACACAAGCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 2700
Db 2641 GCCATCGGCGGTGAGACACAAGCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 2700
QY 2701 GCCTGTGTCACCTGCTCCAGAGCTGCGGCGGGGATTTTCAGAGCGCTTACTCAAGTGT 2760
Db 2701 GCCTGTGTCACCTGCTCCAGAGCTGCGGCGGGGATTTTCAGAGCGCTTACTCAAGTGT 2760
QY 2761 GTGGGCGACGAGCGCGCTGCTGCGCGGACCGAGCTGCAACTTGCACGCAAGCGCCAG 2820
Db 2761 GTGGGCGACGAGCGCGCTGCTGCGCGGACCGAGCTGCAACTTGCACGCAAGCGCCAG 2820
QY 2821 GAGCTGACTTCTCGCTCTGAGCGCGTGTGA 2853
Db 2821 GAGCTGACTTCTCGCTCTGAGCGCGTGTGA 2853

RESULT 5

US-60-216-821-9

; Sequence 9, Application US/60216821

; GENERAL INFORMATION:

; APPLICANT: Deleane, Angelo M.

; APPLICANT: Tang, Y. tom

; APPLICANT: Walsh, Roderick T.

; APPLICANT: Lal, Preeti

; APPLICANT: Khan, Farrah A.

; APPLICANT: Nguyen, Danniell B.

; APPLICANT: Tribouley, Catherine M.

; APPLICANT: Yue, Henry

; APPLICANT: Hafalia, April

; APPLICANT: Patterson, Chandra

; APPLICANT: Lu, Dyung Aina M.

; APPLICANT: Azimzai, Yaida

; APPLICANT: Gandhi, Ameena

; APPLICANT: Baughn, Mariah R.

; APPLICANT: Yao, Monique G.

; APPLICANT: Lee, Ernestine A.

; APPLICANT: Lu, Yan

; APPLICANT: Ramkumar, Jaya

; APPLICANT: Kallick, Deborah A.

; APPLICANT: Au-Young, Janice

; APPLICANT: Xu, Yuming

; TITLE OF INVENTION: PROTEASES

; FILE REFERENCE: FI-0146 P

; CURRENT APPLICATION NUMBER: US/60/216,821

; CURRENT FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PERL Program

; SEQ ID NO 9

; LENGTH: 2930

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: 7473089CB1

US-60-216-821-9

Query Match 98.8%; Score 2817.6; DB 65; Length 2930;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2846; Conservative 0; Mismatches 4; Indels 6; Gaps 2;

QY 1 ATGCTTCTGCTGGGATCTTAACCTTTCGCGGGGGAACCGCTGAGGCTCTGAG 60
Db 75 ATGCTTCTGCTGGGATCTTAACCTTTCGCGGGGGAACCGCTGAGGCTCTGAG 134
QY 61 CCAGAGCGGGAGGTAGTCTGTTCCCATCCGAGCTGACCGGACATTAACGGCCGCGGTAC 120
Db 135 CCAGAGCGGGAGGTAGTCTGTTCCCATCCGAGCTGACCGGACATTAACGGCCGCGGTAC 194

QY 121 TACTGGGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCAGAGCAIT 180
Db 195 TACTGGGGGTCCCGAGGACTCCGGGATCAGGACTCATTTTTCAGATCAGAGCAIT 254
QY 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGGCTCCGCTTCTCC 240
Db 255 CAGGAGGACTTTTACCTACACCTGACGCGGATGCTCAGTTCTTGGCTCCGCTTCTCC 314
QY 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGCTCACCGGGGCTCTTCAGAGCTCGGACGC 300
Db 315 ACTGAGCATCTGGGCTCCCGCTCCAGGGCTCACCGGGGCTCTTCAGAGCTCGGACGC 374
QY 301 TGTCTTCTATTCTGGGAGCTGAACCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 375 TGTCTTCTATTCTGGGAGCTGAACCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 434
QY 361 GGGGGCTCCCGGGAGCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCCGCTGCC 420
Db 435 GGGGGCTCCCGGGAGCTTTGGCTACCGAGCGCGGAGTATGTCATTAGCCCGCTGCC 494
QY 421 AATGCTAGCGCGCGCGCGAGCGCAACAGCGAGGCGCACACCTTCTCCAGCGCGG 480
Db 495 AATGCTAGCGCGCGCGCGAGCGCAACAGCGAGGCGCACACCTTCTCCAGCGCGG 554
QY 481 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTTCGGGCTG 540
Db 555 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGGTGGCTTCGGGCTG 614
QY 541 AACCCCGCATCTACGGGCTTACAGCCCTTACAGCGCGCGGGCGGGCTTCGGGAG 600
Db 615 AACCCCGCATCTACGGGCTTACAGCCCTTACAGCGCGCGGGCGGGCTTCGGGAG 674
QY 601 AGTCTAGCGCGCGAGCTTGGGCGCGCAAGCGTTTCTGCTGCTATCCCGGGTACGTG 660
Db 675 AGTCTAGCGCGCGAGCTTGGGCGCGCAAGCGTTTCTGCTGCTATCCCGGGTACGTG 734
QY 661 GAGAGCTGGTGGTGGCGGAGCTCAATGTCAGGTTCCAGCGCGGAGCTGGAACAT 720
Db 735 GAGAGCTGGTGGTGGCGGAGCTCAATGTCAGGTTCCAGCGCGGAGCTGGAACAT 794
QY 721 TATCTGCTGACGCTGCTGGCAACGGCGCGAGCTTACCGCCATCCAGAGCTTCACAC 780
Db 795 TATCTGCTGACGCTGCTGGCAACGGCGCGAGCTTACCGCCATCCAGAGCTTCACAC 854
QY 781 CCCATCAACATCGTTGTTGCTGAAGTCTTCTTAGAGATCGTGACTCCGGGCCCCAAG 840
Db 855 CCCATCAACATCGTTGTTGCTGAAGTCTTCTTAGAGATCGTGACTCCGGGCCCCAAG 914
QY 841 GTCACCGCAATGCGGCGCTGACGCTGCGCAACTTCTGCTGCTGGCAGAGAGCTGAAC 900
Db 915 GTCACCGCAATGCGGCGCTGACGCTGCGCAACTTCTGCTGCTGGCAGAGAGCTGAAC 974
QY 901 AAAGTGAGTGACAAGCACCCCGAGTACTGGGAGTGGGCTTCCATAGCCTTCCAGGAGGAC 960
Db 975 AAAGTGAGTGACAAGCACCCCGAGTACTGGGAGTGGGCTTCCATAGCCTTCCAGGAGGAC 1034
QY 961 CTGTGTGGAGCCACACCTGTGACACCTGGGATGGTGGTACCATTGTGTGAC 1020
Db 1035 CTGTGTGGAGCCACACCTGTGACACCTGGGATGGTGGTACCATTGTGTGAC 1094
QY 1021 CCCAAGAGAGCTGCTCTGCTCATTTAGAGGAGTGGGCTTCCATAGCCTTCCAGGAGG 1080
Db 1095 CCCAAGAGAGCTGCTCTGCTCATTTAGAGGAGTGGGCTTCCATAGCCTTCCAGGAGG 1154
QY 1081 CAGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1155 CAGAGCTGGGCGACGTGTTCAACATGCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1214
QY 1141 TTTGGGAAGCTCCGAGCAACACCATGATGTCCTCCCGACCTCATCCAGATCGACCGTGC 1200
Db 1215 TTTGGGAAGCTCCGAGCAACACCATGATGTCCTCCCGACCTCATCCAGATCGACCGTGC 1274

QY 1201 AACCCCTGGTCAGCTGAGTGTGCTGCTATCATACCGACTTCTTGAGACGCGGACGGT 1260
DB 1275 AACCCCTGGTCAGCTGAGTGTGCTGCTATCATACCGACTTCTTGAGACGCGGACGGT 1334
QY 1261 GACTGCTCTGAGCAACACCGACAGCCATCTCCCTGCCCGGAGATCTCCCGGGGGCC 1320
DB 1335 GACTGCTCTGAGCAACACCGACAGCCATCTCCCTGCCCGGAGATCTCCCGGGGGCC 1394
QY 1321 AGCTACACCTGAGCAGCAGTGGAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCCT 1380
DB 1395 AGCTACACCTGAGCAGCAGTGGAGCTGGCTTTTGGCGTGGCTCCAAAGCCCTGTCCT 1454
QY 1381 TACATGACGACTGACCAAGCTGTGTGACCGGGAAGGCCAAGGACACAGTGGTGTG 1440
DB 1455 TACATGACGACTGACCAAGCTGTGTGACCGGGAAGGCCAAGGACACAGTGGTGTG 1514
QY 1441 CAGACCGGCACTTCCCTGGGCGGATGGCAGCAGCTGTGGCGAGGGCAAGCTTGGCTC 1500
DB 1515 CAGACCGGCACTTCCCTGGGCGGATGGCAGCAGCTGTGGCGAGGGCAAGCTTGGCTC 1574
QY 1501 AAGGGGCTGCTGGAGAGACACAACCTCAACAGACAGGGTGGATGGTTCCTGGGCC 1560
DB 1575 AAGGGGCTGCTGGAGAGACACAACCTCAACAGACAGGGTGGATGGTTCCTGGGCC 1634
QY 1561 AATGGGATCCCTATGGGCGCTGCTGCGGCACATGTGGTGGGGCGTGCAGCTGGCCAGG 1620
DB 1635 AATGGGATCCCTATGGGCGCTGCTGCGGCACATGTGGTGGGGCGTGCAGCTGGCCAGG 1694
QY 1621 AGGAGTGACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTGGGGTG 1680
DB 1695 AGGAGTGACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTGGGGTG 1754
QY 1681 AATACCGATCTGCAATCTGAGCGCTGCCCGCAGCTCAGCTCCCGGAAGAGCTCCGG 1740
DB 1755 AATACCGATCTGCAATCTGAGCGCTGCCCGCAGCTCAGCTCCCGGAAGAGCTCCGG 1814
QY 1741 GAGGACGAGTGTGAGGCTTTCAACGGCTTCAACACAGCACCAACCGGCTCACCTCGCC 1800
DB 1815 GAGGACGAGTGTGAGGCTTTCAACGGCTTCAACACAGCACCAACCGGCTCACCTCGCC 1874
QY 1801 GTGGATGGTGGCCAGTACTCCGGCTGTCTCCCGGGACAGTGCAGCTCATCTGC 1860
DB 1875 GTGGATGGTGGCCAGTACTCCGGCTGTCTCCCGGGACAGTGCAGCTCATCTGC 1934
QY 1861 CGAGCAATGGCACTGGCTTCTATGTGTGGCACCCAA---GGTGGTGGAGGGCAGC 1917
DB 1935 CGAGCAATGGCACTGGCTTCTATGTGTGGCACCCAAAGGTGTGTGGAGGGCAGC 1994
QY 1918 CTGTGCTCTCTGACTGCTGACCTCGCTGTGTGTCAGAGGCAAGTGCATCAAGCTGGCTGT 1977
DB 1995 CTGTGCTCTCTGACTGCTGACCTCGCTGTGTGTCAGAGGCAAGTGCATCAAGCTGGCTGT 2054
QY 1978 GATGGGAACCTGGCTTCCAGAGAGATTCGACAGTGTGGGGTGTGGGGGAGACAA 2037
DB 2055 GATGGGAACCTGGCTTCCAGAGAGATTCGACAGTGTGGGGTGTGGGGGAGACAA 2114
QY 2038 AAGAGTGCAGAGAGTGTGACTGGACTTTCACCAAGCCCATTCATGGCTTCAATTCGTG 2097
DB 2115 AAGAGTGCAGAGAGTGTGACTGGACTTTCACCAAGCCCATTCATGGCTTCAATTCGTG 2174
QY 2098 GTGGCATCCCGCAGGCGCTTCAAGCATCGACATCCCGCAGCGGTTCAGAGGGCTG 2157
DB 2175 GTGGCATCCCGCAGGCGCTTCAAGCATCGACATCCCGCAGCGGTTCAGAGGGCTG 2234
QY 2158 ATCGGGGATGACAACCTGCTGCTGAGAGACAGCCAGGCAAGTACTGCTCAACGGG 2217
DB 2235 ATCGGGGATGACAACCTGCTGCTGAGAGACAGCCAGGCAAGTACTGCTCAACGGG 2294
QY 2218 CATTTGTGTGTGCGCGGTGAGCGGGACCTGTGTGTGAGAGGGCAGTCTGCTGCGGTAC 2277
DB 2295 CATTTGTGTGTGCGCGGTGAGCGGGACCTGGTGGTGAAGGGCAGTCTGCTGCGGTAC 2354
QY 2278 AGCGGCACGGCACAGCGGTGGAGAGCCTTGACGGCTTCCCGGCCCATCTTGGAGCGCTG 2337

DB 2355 AGCGGCACGGGCACAGCGGTGGAGAGCCTTGACGGCTTCCCGGCCCATCTCTGGAGCGCTG 2414
QY 2338 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCGGGTCCGCTACTCTCTTCTAT 2397
DB 2415 ACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCGGGTCCGCTACTCTTCTAT 2474
QY 2398 CTGCCAAAGACCTTCGGGAGGACAAAGTCTCTCATCCCAAGGACCCCGG---GGACCC 2454
DB 2475 CTGCCAAAGACCTTCGGGAGGACAAAGTCTCTCATCCCGCCGACCCCGGGGAGAGCC 2534
QY 2455 TCTGTCTTGCACACAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGACAGG 2514
DB 2535 TCTGTCTTGCACACAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCAGCGGACGACAGG 2594
QY 2515 CCCCTGACGCTGGTGGCTGGCAGCTGGGGCGCTGCTCCGGAGCTGGCGGAGTGGC 2574
DB 2595 CCCCTGACGCTGGTGGCTGGCAGCTGGGGCGCTGCTCCGGAGCTGGCGGAGTGGC 2654
QY 2575 CTGCAGAGCGGGCTGGAGTGTGGGGCTCCCGCGGAGCGACGCTCCCTGCTGT 2634
DB 2655 CTGCAGAGCGGGCTGGAGTGTGGGGCTCCCGCGGAGCGACGCTCCCTGCTGT 2714
QY 2635 GATGACGCCCATTCGGCCGCTGGAGACACAAGCCTTCGGGGAGCCCTGCCCGAGTGGGAG 2694
DB 2715 GATGACGCCCATTCGGCCGCTGGAGACACAAGCCTTCGGGGAGCCCTGCCCGAGTGGGAG 2774
QY 2695 CTCAGGCTGTGTACCTCTCCAAAGAGCTGGCGCGGGGATTCAGAGCGCTCCTC 2754
DB 2775 CTCAGGCTGTGTACCTCTCCAAAGAGCTGGCGCGGGGATTCAGAGCGCTCCTC 2834
QY 2755 AAGTGTGTGGGCGACGAGCGCGCTGTGTGGCGGGGAGTTCAGAGCGCTCCTC 2814
DB 2835 AAGTGTGTGGGCGACGAGCGCGCTGTGTGGCGGGGAGTTCAGAGCGCTCCTC 2894
QY 2815 CCCAGGAGCTGGACTTCTGTCTGTAGGGCGGTG 2850
DB 2895 CCCAGGAGCTGGACTTCTGTCTGTAGGGCGGTG 2930

RESULT 6

US-10-093-463-27

; Sequence 27, Application US/10093463

; GENERAL INFORMATION:

; APPLICANT: Padigar, Muralidhara

; APPLICANT: Shenoy, Suresh

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Gusev, Vladimir

; APPLICANT: Pochart, Pascal

; APPLICANT: Zhong, Mei

; APPLICANT: Rastelli, Luca

; APPLICANT: Mezes, Peter

; APPLICANT: Smithson, Glennda

; APPLICANT: Guo, Xiaojia

; APPLICANT: Gerlach, Valerie

; APPLICANT: Casman, Stacie

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Tchernev, Velizar

; APPLICANT: Gangolli, Esha

; APPLICANT: Vernet, Corine

; APPLICANT: Pena, Carol

; APPLICANT: Burgess, Catherine

; APPLICANT: Liu, Xiaohong

; APPLICANT: Spytek, Kimberly

; APPLICANT: Gorman, Linda

; APPLICANT: Spaderna, Steven

; APPLICANT: Voss, Edward

; APPLICANT: Malyankar, Uriel

; APPLICANT: Anderson, David

; APPLICANT: Patturajan, Meera

; APPLICANT: Miller, Charles

APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids, and Methods of Use.
FILE REFERENCE: 21402-290A (Cura 590AT)
CURRENT APPLICATION NUMBER: US/10/093,463
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 27
LENGTH: 2804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (25)..(2797)
US-10-093-463-27

Query Match 91.88; Score 2619,2; DB 39; Length 2804;
Best Local Similarity 96.28; Pred. No. 0;
Matches 2753; Conservative 0; Mismatches 13; Indels 96; Gaps 3;
Qy 1 ATGCTTCTGCTGGGATCTAACTGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
Db 25 ATGCTTCTGCTGGGATCTAACTGCTTTCGCGGGGGAACCGCTGGAGGCTCTGAG 84
Qy 61 CCAGAGGGGAGGTAGTCTCCATCGGACTGGACCGGACATTAACGGCCCGGCTAC 120
Db 85 CCAGAGGGGAGGTAGTCTCCATCGGACTGGACCGGACATTAACGGCCCGGCTAC 144
Qy 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCACAGCATTT 180
Db 145 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGCTCATTTTTCAGATCACAGCATTT 204
Qy 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTTGGTCCCGGCTTCTCC 240
Db 205 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTTCTTTGGTCCCGGCTTCTCC 264

Qy 241 ACTGAGCATCTGGCGTCCCTCCAGGGGCTACCGGGGGCTCTTCAGACCTTCGAGCGC 300
Db 265 ACTGAGCATCTGGCGTCCCTCCAGGGGCTACCGGGGGCTCTTCAGACCTTCGAGCGC 324
Qy 301 TGCCTTCTATTTGGGACGTGAAGCGGACCGGAGCTGCTGCTGCTGTGAGCCTGTC 360
Db 325 TGCCTTCTATTTGGGACGTGAAGCGGAGCGGAGCTGCTGCTGCTGTGAGCCTGTC 384
Qy 361 GGGGGCTCCCGGAGCCTTTGGCTACCGAGCGCCGAGTATGTCAATTAGCCCGCTGCC 420
Db 385 GGGGGCTCCCGGAGCCTTTGGCTACCGAGCGCCGAGTATGTCAATTAGCCCGCTGCC 444
Qy 421 AATGCTAGCGCGCGGGCGGCGAGCAACAGCAGGGCGGACACCTTCTCCAGCGCGG 480
Db 445 AATGCTAGCGCGCGGGCGGCGAGCAACAGCAGGGCGGACACCTTCTCCAGCGCGG 504
Qy 481 GGTGTTCCGGGCGGGCCTTCGGGAGACCCACCTCTCGCTCGGGGGTGGCTCCGGGTGG 540
Db 505 GGTGTTCCGGGCGGGCCTTCGGGAGACCCACCTCTCGCTCGGGGGTGGCTCCGGGTGG 564
Qy 541 AACCCCGCATCTACGGGCGCTGGACCTTACAAAGCGGCGGGGGGCTTCGGGGAG 600
Db 565 AACCCCGCATCTACGGGCGCTGGACCTTACAAAGCGGCGGGGGGCTTCGGGGAG 624
Qy 601 AGTCGTAGCGCGCGAGCTCTGGGCGGCGCAAGCGTTCGTGCTATCCCGCGTACGTG 660
Db 625 AGTCGTAGCGCGCGAGCTCTGGGCGGCGCAAGCGTTCGTGCTATCCCGCGTACGTG 684
Qy 661 GAGACGCTGGTGGTGGCGGAGAGTCAATGGTCAAGTTCCACGGCGGAGACTTGAACAT 720
Db 685 GAGACGCTGGTGGTGGCGGAGAGTCAATGGTCAAGTTCCACGGCGGAGACTTGAACAT 744
Qy 721 TATCTGTGAGCTGCTGGCAACGGCGGCGGAGCTCTACCGGCATCCAGCATCTCAAC 780
Db 745 TATCTGTGAGCTGCTGGCAACGGCGGCGGAGCTCTACCGGCATCCAGCATCTCAAC 804
Qy 781 CCCATCAACATCGTTGTGTCAGGTGCTTCTTAGAGATCGTGACTCCGGGCGCAAG 840
Db 805 CCCATCAACATCGTTGTGTCAGGTGCTTCTTAGAGATCGTGACTCCGGGCGCAAG 864
Qy 841 GTACCGGCAATGCGGCGCTGACGCTGGCGCAACTTCTGTGCTGGCAAGAGCTGAAC 900
Db 865 GTACCGGCAATGCGGCGCTGACGCTGGCGCAACTTCTGTGCTGGCAAGAGCTGAAC 924
Qy 901 AAAGTGAGTGACAGCACCCCGAGTCTGGGACACTGCCATCTCTCCACAGGAGSAC 960
Db 925 AAAGTGAGTGACAGCACCCCGAGTCTGGGACACTGCCATCTCTCTCCACAGGAGSAC 984
Qy 961 CTGTGTGAGCCACCACTGTGACACCTTGGCATGGCTGATGGGTACCAATGTGAGC 1020
Db 985 CTGTGTGAGCCACCACTGTGACACCTTGGCATGGCTGATGGGTACCAATGTGAGC 1044
Qy 1021 CCCAAGAGAGTGTCTGTCTATTGAGGACGATGGGCTTCCATAGCCTTCCACTGCC 1080
Db 1045 CCCAAGAGAGTGTCTGTCTATTGAGGACGATGGGCTTCCATAGCCTTCCACTGCC 1104
Qy 1081 CAGGAGCTGGGCGGAGCTGTTCAACATGCCCATGACCAATGTGAAAGTCTGTGAGGAGTG 1140
Db 1105 CAGGAGCTGGGCGGAGCTGTTCAACATGCCCATGACCAATGTGAAAGTCTGTGAGGAGTG 1164
Qy 1141 TTTGGAAGCTCCGAGCGCAACCAACATGATGTCCCGAGCCCTCATCCAGATCGACCGTGC 1200
Db 1165 TTTGGAAGCTCCGAGCGCAACCAACATGATGTCCCGAGCCCTCATCCAGATCGACCGTGC 1224
Qy 1201 AACCCCTGTGTAGCTGTGAGTGTGCTGCTATCATCCGACTTCTCCGAGCGGCGGACGCT 1260
Db 1225 AACCCCTGTGTAGCTGTGAGTGTGCTGCTATCATCCGACTTCTCCGAGCGGCGGACGCT 1284
Qy 1261 GACTGCTCTCTGAGCAACCCAGAGCGGCAATCTCCCTCCCGAGGATCTGCGGGCGGCG 1320
Db 1285 GACTGCTCTCTGAGCAACCCAGAGCGGCAATCTCCCTCCCGAGGATCTGCGGGCGGCG 1344
Qy 1321 AGCTACACCTTGAGCGAGCGAGTGTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCTCT 1380

301	Db	TGCTTTCTATTCCGGGTATGTGAACCGCGAGCCAGACTTCCTTTGCTGTTGAGCCTATATGC	360
361	QY	GGGGGGCTCCGCGGAGCCTTTGGCTACCGAGGGCCGAGTATGCTATTATACCGCGCTGCC	420
361	Db		
361	Db	GGGGTCTCCGCGGAGCCTTTGGCTACCGAGTGCAGGATATGCTATTAGCCTCTGCC	420
421	QY	AATGCTAGCGGCGCGGCGCGAGCGCAACAGCAGCGGCCACACCTTCTCCAGCGCCGG	480
421	Db		
421	Db	AACACAGCGCGCCAGAGAGCGCGCTCACAGCAGCGGCACACCTTCTCCAGCGCCGG	480
481	QY	GGTGTTCGGGCGGGCCTTCGCGAGACCCACCTCTCGCTCGGGGTGGCCTCGGGCTGG	540
481	Db		
481	Db	GGTGCTCTGTAGAGCCCTTCGCGAGACCCACCTCTCGTTGGGGGTGGCCTCGGGCTGG	540
541	QY	AACCCCGCCATCCTTACGGGCGCTTGAACCTTTACAAGCGCGCGGGCGGGGCTTCGGGGAG	600
541	Db		
541	Db	AACCCCGCCATCCTTACGGGCTCTGGACCTTTATAAGCACAGCGCGAGCGCGCGGGGGAG	600
601	QY	AGTCGTAGCGCGCGAGCTCTGGGCGGCCAAGGTTTCGTGTCTATCCCGCGGTACGTG	660
601	Db		
601	Db	AGCCACACCGCGCGAGGTCGTGGGCGGCCAAGCGCTTCGTGTCTATACACAGGTACGTG	660
661	QY	GAGACCTGGTGTTCGCGGACGAGTCAATGTGTCAAGTTCACGCGCGGACCTTGGAACT	720
661	Db		
661	Db	GAGACACTGGTGTTCGCGGACGAGTCAATGGTCAAGTTCACGCGCGGATTTGGAACT	720
721	QY	TATCTGCTGACGCTGCTGGCAACGCGCGGACTCTACGCCATCCACGATCCTCAAC	780
721	Db		
721	Db	TATCTACTGACGCTGCTGGCCACGCGCGGAGCTCTACGCCACCCACGATCCTCAAC	780
781	QY	CCCATCAACATCGTTGTGGTCAAGTGCTGCTTCTTAGAGATCGTGACTCCGGGCCAAG	840
781	Db		
781	Db	CCTATCAACATCGTGTGGTCAAGTGCTACTCTTAGAGATCGTGACTCGGCGCCAAG	840
841	QY	GTACCGGCAATGCGGCCCTGACGCTGGCAACTCTCTGCTGGCAGACAGCAAGCTGAAC	900
841	Db		
841	Db	GTACAGGCAACGCGCCCTGACTCTGGCGCAACTTCTGTGCTGGCAGAAAGCTGAAC	900
901	QY	AAAGTGAGTGACAAGCACCCCGAGTACTGGGACATGCGCATCTCTTCCACGSCAGGAC	960
901	Db		
901	Db	AAAGTGAGCGATAAGCACCCCTGAGTACTGGGACACAGCCATCTCTTCCACGACAGGAC	960
961	QY	CTGTGTGGAGCCACCACTGTGACACCTTGGGCATGGCTGATGTGGTACCATGTGTAC	1020
961	Db		
961	Db	CTATGGGGGCTACCACTGTGACACCTTGGGTATGGGTGATGTGGGACCATGTGTGAC	1020
1021	QY	CCAAAGAGAGCTGCTCTGCTATTGAGGACGATGGGCTTCCATCAGCCTTACCACTGCC	1080
1021	Db		
1021	Db	CCCAAGAGAGCTGCTCTGTCAATTGAGGACGATGGCTTCCATCAGCCTTACCACTGCC	1080
1081	QY	CAGAGGTGGGCCACGTGTTCAACATGCCCCATGACATGTGAAGTCTGTGAGGAGTG	1140
1081	Db		
1081	Db	CATGAGCTGGGCCATGTGTTCAACATGCCCCATGACAAGTGAAGGTGTGTGAGGAGTG	1140
1141	QY	TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCTCATCCAGATGCACCGTGCC	1200
1141	Db		
1141	Db	TTTGGGAAGCTCAGGCCAACCAACCATGATGTCTCGACCTCATCCAGATGCACCGTGCC	1200
1201	QY	AACCCCTGGTCAGCCTGCATGCCATCATCACCGACTTCCTGGACAGCGGCGACGGT	1260
1201	Db		
1201	Db	AACCCCTGGTCAGCCTGCAGCGTGCCATTATTACTGACTTCTGGATAGTGGGCATGGT	1260
1261	QY	GACTGCTCTCTGGNACCAACCCACAGACCCATCTCCCTGCCGAGGATCTCCGGGGGCC	1320
1261	Db		
1261	Db	GACTGCTCTCTGGACACGCCCCAAGCCCATCAACCTTGCCTGAGGACTCTCCGGGCACA	1320
1321	QY	AGCTTACACCTTGAGCCAGCAGTGCAGCTGGCTTTTGGCGTGGGCTCCAAGCCCTGTCT	1380
1321	Db		
1321	Db	AGCTTACAGTTTGACCAAGCAGTGCAGCTGGCTTTTGGGTGGGCTCTAAGCCTTGCCCA	1380
1381	QY	TACATGCAAGTACTGCACCAAGCTGTGGTGACCGGGGAAGGCCAAGGCACATGGTGTGC	1440
1381	Db		
1381	Db	TATATGCAAGTACTGTACAAAGCTGTGGTGACTTGCAGGCGCAAGGGCAATGGTGTGC	1440

Qy	1441	CAGACCCGCACTTCCCCTGGCGCATGGCACAGCTGTGGCGAGGGCAAGCTCTGGCTC	1500
Db	1441	CAAGTCTGCCACTTCCCCTGGCGCATGGCACAGCTGTGGCGAAGGCAAGTCTTGGCTC	1500
Qy	1501	AAAGGGGCTGGTGGAGAGACACAACCTCAACAAGCACAGGTGGATGGTTCTCTGGGCC	1560
Db	1501	AAGGAGGCGTGGTGGAGAGACACAACCAACAAGTACCGGGTGGACGGCTCTTGGGCC	1560
Qy	1561	AAATGGGATCCCTATGGCCCTGCTCGGCGACATGTGTGGGGCGTGCAGCTGGCCAGG	1620
Db	1561	AAATGGGAGCCCTACGGTCTCTCCGACCTCGCGTGGGGCGTGCAGCTGGCCCGG	1620
Qy	1621	AGGAGTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTGCGAGGGAGTGAGGGTG	1680
Db	1621	AGGAGTGCAGCAACCCCTACCCCTGCCAACGGTGGAAATACTGCGAGGGAGTAAGATG	1680
Qy	1681	AAATACCGATCTGCAATCTGAGCGCTGCCCCAGCTCAGCCTCCGGAAAGAGTCTCCGG	1740
Db	1681	AAATACCGATCTTGCACTTGAGGCCCTGCCCCAGCTCAGCCTCTGGCAAGAGTCTCCGG	1740
Qy	1741	GAGGAGCACTGTGAGGCTTTCAACGGCTACAACACAGCACCAACCGGCTCACTCTGCC	1800
Db	1741	GAAGAGCAATGTGAAGCTTTCAATGGCTACAACACAGCACCAACCGGCTCACTTTAGCT	1800
Qy	1801	GTGGCATGGTGGCCCAAGTACTCGGCGGTGTCTCCCGGGACAAGTGCAAAGCTCATCTGC	1860
Db	1801	GTGGCATGGGTACCCCAAGTACTCAGCGGTGTCAACACGTGACAAAGTGAAGCTCATCTGC	1860
Qy	1861	CGAGCAATGGCACTGGCTACTTCTATGTGTGTGGCACCAAGTGTGGACGGCAGCTGTG	1920
Db	1861	CGAGCAATGGGACTGGCTACTTCTATGTACTAGCACCTTAAGTGTGTGATGTACGGCTG	1920
Qy	1921	TGCTCTCTGTACTCCACCTCCGCTGTGTGTCCAGGCAAGTGCATCAAGGCTGGCTGTGAT	1980
Db	1921	TGTACTCTGTACTCCACCTTCGGTGTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGCGAC	1980
Qy	1981	GGGAACCTGGGCTTCCAAGAGAGATTGCAAGTGTGGGTGTGTGGGGAGACAAATAAG	2040
Db	1981	GGGAATCTGGGCTCCAAGAAGAAATTTGACAAATGTGGTGTGTGGTGGAGACAAATAAG	2040
Qy	2041	AGCTGCAAGAGTGTACTGGACTTCTCACCACGGCCATGCATGGCTACAAATTTCTGGTGTG	2100
Db	2041	AGCTGTAAGAGGTGTACAGGACTTCTCACCAGCCATGTGACGGCTACAAATTTGTAGTGTG	2100
Qy	2101	GCCATCCCCGACGGCGCTCAAGCATCGACATCCGCCAGCGGTTTACAAAGGGCTGATC	2160
Db	2101	GCCATCCCTCCCGCGCCTCCAGCATTTGATCCGTGAGCTGGTTTACAAAGGGGCTCAAT	2160
Qy	2161	GGGATGACAACCTACCTGGCTCTGAAGACAGCCCAAGGCAAGTACCTGCTCAACGGGCAT	2220
Db	2161	GGAGATGACAACCTACCTGGCTTTGAAGACAGCCCAAGGCAAAATACCTGCTCAATGGGCAC	2220
Qy	2221	TTCTGTGTGTCTGGCGGTGGAGCGGACCTGGTGGTGAAGGGCAAGTGTCTGCGGTACAGC	2280
Db	2221	TTTGTGTATCCGCTGTAGACGGGACCTGGTGGTAAAGGCAAGTGTGTACGCTATAGT	2280
Qy	2281	GGCACGGGCACAGGCTGTGAGAGCTGTGAGGCTTCCGGGCCCATCTCTGGAGCGCTGACC	2340
Db	2281	GGCACTGGCACTGCAGTGTGAGAGGCTTCAAGGCTTCTGACCCATCTCTGGAGCCACTCAC	2340
Qy	2341	GTGAGGTCTCTCCGTGGGAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTGTG	2400
Db	2341	GTGAGGTCTCTGTCTGGTGGGAGATGACACCGCCCGCTGTGGGTATTCTTCTACCTGTG	2400
Qy	2401	CCCAAGAGGCTCTGGGAGGCAAGTCTCTCATCCCCAAGACCCCCGGGACCCCTCTGTCT	2460
Db	2401	CCCAAGAACCTCTGGGAGGACAAGTCTCACTCGCCCCAAGGACCCCTCGAGGTCTCCCGGTG	2460
Qy	2461	TTGCACAACAGGCTCTACGCTCTCCAACAGGTGTGAGCAGCGGACGACAGGCCCTCT	2520
Db	2461	TTGGCAACAGTGTCTCAGCCTTCCACCAAGTATGAGCAGCGCAATATCAACGACCTCTCT	2520

```
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGCGTCTCCGCGAGCTGGCGAGTGGCGCTGCAG 2580
Db 2521 GCGCGCTGGGTGGCAGCAGCTGGGACCTTGTCTAGTGAAGTGGGAGTGGGCTACAG 2580
QY 2581 AAGCGGGGGTGGAGTGTCTGGGGTTCGCCGGGCGACGCGACGGTCCCTGCTGTGTGATGCA 2640
Db 2581 AAGCGAGAGTGGAGTGTCTGGGGTTCGCCGGGCGACGCGACGGTTCGTGTGTGATGTA 2640
QY 2641 GCCCATCGCGCGTGGAGACACAGCTTGGCGGGGAGCGCTCCGCCACCTGGAGAGCTCAGC 2700
Db 2641 GACCATCGCGCGTGGAGAGCGAGCTTGGCGGGGAGCGCTCCGCCACCTGGAGAGCTCAGC 2700
QY 2701 GCCTGGTCAACCTGTCTCAAGAGCTGGCGCGGGGATTTACAGAGCGCTCACTCAAGTGT 2760
Db 2701 AACTGGTGGCGCTGTCTCAAGAGCTGGCGCGGGGATTTAAGAGCGCTCAAGTGT 2760
QY 2761 GTGGGCCACGAGGCGCGCTGTCTGGCGGGGACGAGTGAACCTTGGACCGCAAGCCCGCAG 2820
Db 2761 GTGGGCCACGAGGCGCGCTGTCTGGCGGGGACGAGTGAACCTTACGCGCAAGCCCGCAG 2820
QY 2821 GAGCTGGAGTCTCGCTCGCTCAGGCGCTGTGA 2853
Db 2821 GAATTAGACTTCTGCTCTTGAGACCTCGCTGA 2853

RESULT 8
US-10-163-316-3
; Sequence 3, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP101-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2469)
US-10-163-316-3

Query Match 80.5%; Score 2298; DB 41; Length 2469;
Best Local Similarity 97.0%; Pred. No 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;

QY 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTTCGCCGGGCGAACCCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTTCGCCGGGCGAACCCTGGAGGCTCTGAG 60
QY 61 CCAGAGGGGAGTGTCTGTTCCATCCGACTGGACCCGAGACATTAACCGGCGCGCTAC 120
Db 61 CCAGAGGGGAGTGTCTGTTCCATCCGACTGGACCCGAGACATTAACCGGCGCGCTAC 120
QY 121 TACTGGCGGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCAGACATTT 180
Db 121 TACTGGCGGGGTCCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCAGACATTT 180
QY 181 CAGGAGACTTTTACCTACACTGACCGCGGATGTCTAGTTCTTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGACTTTTACCTACACTGACCGCGGATGTCTAGTTCTTGGCTCCCGCTTCTCC 240
QY 241 ACTGACACTTGGCGCTCCCGCTCCAGGGGCTCACCAGGGGCTCTTTCAGACCTGGAGC 300
Db 241 ACTGACACTTGGCGCTCCCGCTCCAGGGGCTCACCAGGGGCTCTTTCAGACCTGGAGC 300
QY 301 TGCTTCTATTCTGGGACGTGAACCGCGAGCGGACTTCGCTCTGAGCCTGTGC 360
Db 301 TGCTTCTATTCTGGGACGTGAACCGCGAGCGGACTTCGCTCTGAGCCTGTGC 360
QY 361 GGGGGGCTCCGCGAGCCTTTTGGCTACCGAGGCGCGAGTATGTCTATTAGCCCGCTGCCC 420
Db 361 GGGGGGCTCCGCGAGCCTTTTGGCTACCGAGGCGCGAGTATGTCTATTAGCCCGCTGCCC 420
QY 421 AATGTAGCGCGCGCGCGCGAGCGCAACAGCAGCGGCGCACACCTTCTCCAGCGCGG 480
Db 421 AATGTAGCGCGCGCGCGCGAGCGCAACAGCAGCGGCGCACACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCGGGCGCGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGSCCTCGGGCTG 540
Db 481 GGTGTTCGGGCGCGGCTTCCGGAGACCCACCTCTCGCTGGGGGTGSCCTCGGGCTG 540
QY 541 AACCCCGCATCTACGGCGCTTGGACCTTACAAAGCGCGGCGGCGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTACGGCGCTTGGACCTTACAAAGCGCGGCGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGCGAGTCTGGCGCGCAAGCGTTTCTGTCTATCCCGGCTACGTG 660
Db 601 AGTCGTAGCGCGCGAGTCTGGCGCGCAAGCGTTTCTGTCTATCCCGGCTACGTG 660
QY 661 GAGACCTGTGTCTCGCGGAGTCAATGGTCAAGTTCACGGCGCGGCTTGGAAAT 720
Db 661 GAGACCTGTGTCTCGCGGAGTCAATGGTCAAGTTCACGGCGCGGCTTGGAAAT 720
QY 721 TATCTCTGACGCTGTCTGCAAGCGCGGCGGCTTACCGGCATCCAGCATCTCAAC 780
Db 721 TATCTCTGACGCTGTCTGCAAGCGCGGCGGCTTACCGGCATCCAGCATCTCAAC 780
QY 781 CCCATCAACATCGTTGTGTCTCAAGTGTCTTCTAGAGATCGTACTCCGGGCGGCGG 840
Db 781 CCCATCAACATCGTTGTGTCTCAAGTGTCTTCTAGAGATCGTACTCCGGGCGGCGG 840
QY 841 GTCACGGCAATGCGGCGCTGACCGTTCGCTTCTTGTCTGCGCAAGAGTGAAC 900
Db 841 GTCACGGCAATGCGGCGCTGACCGTTCGCTTCTTGTCTGCGCAAGAGTGAAC 900
QY 901 AAAGTGAGTGAACGACCGCGAGTCTGGGACCTGCGGCTTCTTGTCTGCGCAAGAGTGAAC 960
Db 901 AAAGTGAGTGAACGACCGCGAGTCTGGGACCTGCGGCTTCTTGTCTGCGCAAGAGTGAAC 960
QY 961 CTGTGTGGAGCGACCGCTGTGACACCTTGGCGGCTGATGTGGTACCATGTGTGAC 1020
Db 961 CTGTGTGGAGCGACCGCTGTGACACCTTGGCGGCTGATGTGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGTGTCTGTCTCAATGAGGAGTGGCTTCCATCAGCCTTCAACACTGCC 1080
Db 1021 CCCAAGAGAGTGTCTGTCTCAATGAGGAGTGGCTTCCATCAGCCTTCAACACTGCC 1080
QY 1081 CACGAGCTGGGCGCGTGTTCACCATGCCCATGACATGTGAAGTCTGTGAGAGGTG 1140
Db 1081 CACGAGCTGGGCGCGTGTTCACCATGCCCATGACATGTGAAGTCTGTGAGAGGTG 1140
QY 1141 TTTGGGAAGTCTCCGAGCGCAACACATGATGTCCCGAGCGCTCATCCAGATCGACTGCC 1200
Db 1141 TTTGGGAAGTCTCCGAGCGCAACACATGATGTCCCGAGCGCTCATCCAGATCGACTGCC 1200
QY 1201 AACCCCTGTGAGCTGTGCTGCCATCATACCGGCTTCTTGGAGCGGGGACGGT 1260
Db 1201 AACCCCTGTGAGCTGTGCTGCCATCATACCGGCTTCTTGGAGCGGGGACGGT 1260
QY 1261 GACTGCTTCTTGGAGCGCAACCGAGCGGCTTCTCCCTGCCCGGAGGATCTCCCGGCGG 1320
Db 1261 GACTGCTTCTTGGAGCGCAACCGAGCGGCTTCTCCCTGCCCGGAGGATCTCCCGGCGG 1320
QY 1321 AGCTACACCTTGGAGCGAGTGGGCTTTTGGGCTGGGCTTCAAGCCCTGTCT 1380
Db 1321 AGCTACACCTTGGAGCGAGTGGGCTTTTGGGCTGGGCTTCAAGCCCTGTCT 1380
QY 1381 TACATGAGTACTGACCAAGCTGTGTGACCGGAGCGCAAGGACAGATGTGTGC 1440
Db 1381 TACATGAGTACTGACCAAGCTGTGTGACCGGAGCGCAAGGACAGATGTGTGC 1440
```



```
Db 1381 TACATGAGTACTGACCAAGCTGTGGTGCACCGGGAAGGCAAGACAGATGGTGTGC 1440
Qy 1441 CAGACCGGCACATTCCTCGGCGCGATGGACACAGCTGTGGGAGGCGAAGCTCTGCCTC 1500
Db 1441 CAGACCGGCACATTCCTCGGCGCGATGGACACAGCTGTGGGAGGCGAAGCTCTGCCTC 1500
Qy 1501 AAAGGGGCTCGCTGGAGAGACACACCTCAACAAGACACAG-----1541
Db 1501 AAAGGGGCTCGCTGGAGAGACACACCTCAACAAGACACAGGCTGACATCAT 1560
Qy 1542 -----GGTGGATGGT 1551
Db 1561 TCTCCAAAACAACCTTTATTAAAGCTACCAATATGGCTGCACACTACACAGGTGGATGGT 1620
Qy 1552 TCCTGGGCCAAATGGGATCCTATGGCCCTGTCTCGGCACATGTGTGGGGCGTGCAG 1611
Db 1621 TCCTGGGCCAAATGGGATCCTATGGCCCTGTCTCGGCACATGTGTGGGGCGTGCAG 1680
Qy 1612 CTGGCCAGGAGGAGTGCACCAACCCCTGCCCCCAACGGGGGCAAGTACTGCGAGGA 1671
Db 1681 CTGGCCAGGAGGAGTGCACCAACCCCTGCCCCCAACGGGGGCAAGTACTGCGAGGA 1740
Qy 1672 GTGAGGTGAATATCCGATCCTGGAATCTGGAGCCCTGCCACAGCTCAGCTCCGGAAG 1731
Db 1741 GTGAGGTGAATATCCGATCCTGGAATCTGGAGCCCTGCCACAGCTCAGCTCCGGAAG 1800
Qy 1732 AGCTTCGGGAGGAGCAGTGTAGGCTTTCAACGGCTTACAAACACACACACCAACCGGCTC 1791
Db 1801 AGCTTCGGGAGGAGCAGTGTAGGCTTTCAACGGCTTACAAACACACACACCAACCGGCTC 1860
Qy 1792 ACTCTCGCGGTGGATGGGTGCCAAGTACTCCGGGCTGTCTCCCGGGGACAAAGTCAAG 1851
Db 1861 ACTCTCGCGGTGGATGGGTGCCAAGTACTCCGGGCTGTCTCCCGGGGACAAAGTCAAG 1920
Qy 1852 CTCATCTGCGAGGCAATGACCTGAGCTACTTCTATGTCTGTCACCAAGTGGTGGAC 1911
Db 1921 CTCATCTGCGAGGCAATGACCTGAGCTACTTCTATGTCTGTCACCAAGTGGTGGAC 1980
Qy 1912 GGCAGCTGTCTCTCTGACTCCACCTCCGCTGTCTGTCTCAAGGCAAGTGCATCAAGGCT 1971
Db 1981 GGCAGCTGTCTCTCTGACTCCACCTCCGCTGTCTGTCTCAAGGCAAGTGCATCAAGGCT 2040
Qy 1972 GGCTGTGATGGGAACCTCGGCTCCAAAGAGATTCGACAAAGTGTGGGGTGTGTGGGGGA 2031
Db 2041 GGCTGTGATGGGAACCTCGGCTCCAAAGAGATTCGACAAAGTGTGGGGTGTGTGGGGGA 2100
Qy 2032 GACAAATAGAGCTGCAAGAGGTGACTGGACTCTTCAACCAAGCCCATGCAATGCTACAAT 2091
Db 2101 GACAAATAGAGCTGCAAGAGGTGACTGGACTCTTCAACCAAGCCCATGCAATGCTACAAT 2160
Qy 2092 TTCGTGGTGGCCATCCCGCAGGCGCCTCAAGCATCGACATCCGCCAGCGGTTACAAA 2151
Db 2161 TTCGTGGTGGCCATCCCGCAGGCGCCTCAAGCATCGACATCCGCCAGCGGTTACAAA 2220
Qy 2152 GGGCTGTATCGGGGATGACAACTACTGGCTCTGAAGAACGCCAAGGCAAGTACCTGCTC 2211
Db 2221 GGGCTGTATCGGGGATGACAACTACTGGCTCTGAAGAACGCCAAGGCAAGTACCTGCTC 2280
Qy 2212 AACGGGCAATTCGTGGTGTCTCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTG 2271
Db 2281 AACGGGCAATTCGTGGTGTCTCGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGCTG 2340
Qy 2272 CCGTACAGCGGCACGGGCACAGCGGTGGAGAGCCCTGCAAGGCTTCCCGGCCCATCCTGGAG 2331
Db 2341 CCGTACAGCGGCACGGGCACAGCGGTGGAGAGCCCTGCAAGGCTTCCCGGCCCATCCTGGAG 2400
Qy 2332 CCGCTGACCGGTGGAGGTCTCTCCGTGGGGAGAGATGACACCGCCCGGGTCCG 2385
Db 2401 CCGCTGACCGGTGGAGGTCTCTCCGTGGGGAGAGATGACACCGCCCGGGTCCG 2454
```

RESULT 9
US-60-297-863-3

```
; Sequence 3, Application US/60297863
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses
; FILE OF INVENTION: Therefor
; CURRENT APPLICATION NUMBER: US/60/297,863
; CURRENT FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(2469)
US-60-297-863-3
```

Query Match 80.5%; Score 2298; DB 73; Length 2469;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;

```
Qy 1 ATGCTTCTGCTGGGCATCCTAACCTGGCTTTCGCCGGGCAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGCATCCTAACCTGGCTTTCGCCGGGCAACCGCTGGAGGCTCTGAG 60
Qy 61 CCAGAGCGGGAGTGTAGTGTCCCATCCGACTGGACCCGAGCATTTAAGCGCGCGCTAC 120
Db 61 CCAGAGCGGGAGTGTAGTGTCCCATCCGACTGGACCCGAGCATTTAAGCGCGCGCTAC 120
Qy 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCACAGCATTT 180
Db 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGGACTCATTTTCAGATCACAGCATTT 180
Qy 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGTCTAGTTCTTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACGCGGATGTCTAGTTCTTGGCTCCCGCTTCTCC 240
Qy 241 ACTGAGCATCTGGGCTCCCCCTCCAGGGGCTACCGGGGCTCTTCAGAGCTTGGAGCGC 300
Db 241 ACTGAGCATCTGGGCTCCCCCTCCAGGGGCTACCGGGGCTCTTCAGAGCTTGGAGCGC 300
Qy 301 TGTCTTATCTGGGAGCGTGAACCCGAGCGGACTGTTGCTGCTGCTGTGAGCCTGTGC 360
Db 301 TGTCTTATCTGGGAGCGTGAACCCGAGCGGACTGTTGCTGCTGCTGTGAGCCTGTGC 360
Qy 361 GGGGGCTCCGCGAGCCTTTGGCTACCGAGCGCGGAGTATGTCTATTAGCCCGTGGCC 420
Db 361 GGGGGCTCCGCGAGCCTTTGGCTACCGAGCGCGGAGTATGTCTATTAGCCCGTGGCC 420
Qy 421 AATGCTAGCGCGCGGCGGCGAGCAACAGCCAGGGGCGACACTTCTCCAGCGCGCG 480
Db 421 AATGCTAGCGCGCGGCGGCGAGCAACAGCCAGGGGCGACACTTCTCCAGCGCGCG 480
Qy 481 GGTGTTCCGGGCGGCTTCCGGAGACCCCACTCTCGCTGCGGGTGGCTTGGGCTGG 540
Db 481 GGTGTTCCGGGCGGCTTCCGGAGACCCCACTCTCGCTGCGGGTGGCTTGGGCTGG 540
Qy 541 AACCCCGCCATCCTACGGGCGCTTGAACCCCTTAAAGCCCGCGCGGCGGCTTCGGGAG 600
Db 541 AACCCCGCCATCCTACGGGCGCTTGAACCCCTTAAAGCCCGCGCGGCGGCTTCGGGAG 600
Qy 601 AGTCGTAGCGCGCGAGTCTGGGCGCGCAAGCGTTTGTGTCTATCTCCCGGCTACGTG 660
Db 601 AGTCGTAGCGCGCGAGTCTGGGCGCGCAAGCGTTTGTGTCTATCTCCCGGCTACGTG 660
Qy 661 GAGACGCTGTGTGTCGCGGACGAGTCAATGGTCAAGTTCACGGCGCGGACCTTGAACAT 720
Db 661 GAGACGCTGTGTGTCGCGGACGAGTCAATGGTCAAGTTCACGGCGCGGACCTTGAACAT 720
Qy 721 TATCTGCTGACGCTGTGGCAACGGCGCGGACTCTACCGCATCCCGAGCATCTCTCAAC 780
Db 721 TATCTGCTGACGCTGTGGCAACGGCGCGGACTCTACCGCATCCCGAGCATCTCTCAAC 780
```



```
Db 721 TATCTGCTGACGCTGCTGCAACGGCGGCACTCTACGGCCATCCCAAGCATCTCAAC 780
QY 781 CCCATCAACATCGTTGCTGAAGTGTGCTTCTTAGAGATCGTGACTCCGGGCCAAG 840
Db 781 CCCATCAACATCGTTGCTGAAGTGTGCTTCTTAGAGATCGTGACTCCGGGCCAAG 840
QY 841 GTACCGGCAATCGCGCCCTGACGCTGGCGCAACTTCTGTGCTTGCAGAGAAAGCTGAAC 900
Db 841 GTACCGGCAATCGCGCCCTGACGCTGGCGCAACTTCTGTGCTTGCAGAGAAAGCTGAAC 900
QY 901 AAAGTGAAGTGAACACCCGAGTACTGCGACACTGCCATCCGTCCTCACAGCAGGAC 960
Db 901 AAAGTGAAGTGAACACCCGAGTACTGCGACACTGCCATCCGTCCTCACAGCAGGAC 960
QY 961 CTGTGTGAGACCAACCATCTGTGACACCTGGGCATGGCTGATG TGGTACCATGTGTGAC 1020
Db 961 CTGTGTGAGACCAACCATCTGTGACACCTGGGCATGGCTGATG TGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAAGTGTCTGTCTGATGAGACGATGGCTTCCATAGGCTTCACCACTGCC 1080
Db 1021 CCCAAGAGAAGTGTCTGTCTGATGAGACGATGGCTTCCATAGGCTTCACCACTGCC 1080
QY 1081 CAGGAGCTGGCGACGCTGTTCAACATGCCCATGACATGTAAGTCTGTGAGAGGTG 1140
Db 1081 CAGGAGCTGGCGACGCTGTTCAACATGCCCATGACATGTAAGTCTGTGAGAGGTG 1140
QY 1141 TTTGGGAAGCTCGGAGCCCAACACATGATGTCCCGGACCTCA TCCAGATGGACCTGCC 1200
Db 1141 TTTGGGAAGCTCGGAGCCCAACACATGATGTCCCGGACCTCA TCCAGATGGACCTGCC 1200
QY 1201 AACCCCTGTGACGCTGCACTGCTGCCATCATCACGACTTCC TGGCAGCGGCACGGT 1260
Db 1201 AACCCCTGTGACGCTGCACTGCTGCCATCATCACGACTTCC TGGCAGCGGCACGGT 1260
QY 1261 GACTGCTCTGAGCAACCCAGACGATCTCCCTGCGGAGGATCTGCCGGGCGCC 1320
Db 1261 GACTGCTCTGAGCAACCCAGACGATCTCCCTGCGGAGGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCCCTGAGCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
Db 1321 AGCTACACCCCTGAGCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTCT 1380
QY 1381 TACATGCACTGATGCAACCAAGCTGTGTGACCGGGAAGCCAGGACACAGATGGTGTGC 1440
Db 1381 TACATGCACTGATGCAACCAAGCTGTGTGACCGGGAAGCCAGGACACAGATGGTGTGC 1440
QY 1441 CAGACCGGCACTTCCCTGGGCGGATGSCACAGCTGTGCGAGGGCAAGCTCTGCCTC 1500
Db 1441 CAGACCGGCACTTCCCTGGGCGGATGSCACAGCTGTGCGAGGGCAAGCTCTGCCTC 1500
QY 1501 AAAGGGGCTGCGTGAGAGACACAACTTCAACAGCAGACAG----- 1541
Db 1501 AAAGGGGCTGCGTGAGAGACACAACTTCAACAGCAGCAG----- 1541
QY 1542 -----GCTGATGTT 1551
Db 1542 -----GCTGATGTT 1551
QY 1552 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCACATGTGTGGGGGCGTGCAG 1611
Db 1552 TCCTGGGCCAAATGGGATCCCTATGGCCCTGCTCGCACATGTGTGGGGGCGTGCAG 1611
QY 1621 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGCACATGTGTGGGGGCGTGCAG 1680
Db 1621 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGCACATGTGTGGGGGCGTGCAG 1680
QY 1661 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGCACATGTGTGGGGGCGTGCAG 1740
Db 1661 CTGGCAGGAGGAGTGCACCAACCCACCCCTGCGCACATGTGTGGGGGCGTGCAG 1740
QY 1672 GTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCGGAGGAGGAG 1731
Db 1672 GTGAGGTGAATACCGATCTGCAATCTGGAGCCCTGCGGAGGAGGAG 1731
QY 1732 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACAGCAGCAGCAGCGGCTC 1791
Db 1732 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACAGCAGCAGCAGCGGCTC 1791
QY 1801 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACAGCAGCAGCAGCGGCTC 1860
Db 1801 AGCTTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTCAACAGCAGCAGCAGCGGCTC 1860
```

```
QY 1792 ACTCTGCGCGTGGCATGGTGCCCAAGTACTCCGGCGCTGTCTCCCGGGACAAGTCAAG 1851
Db 1861 ACTCTGCGCGTGGCATGGTGCCCAAGTACTCCGGCGCTGTCTCCCGGGACAAGTCAAG 1920
QY 1852 CTCATCTCCGAGGCAATGGCACTGGCTACTTCTATGTGTGGTGGACCAAGTGGTGGAC 1911
Db 1921 CTCATCTCCGAGGCAATGGCACTGGCTACTTCTATGTGTGGTGGACCAAGTGGTGGAC 1980
QY 1912 GSCAGCGTGTCTCTCTGACTCCACCTCCGCTGTCTTCCAAAGGCAAGTGCATCAAGGCT 1971
Db 1981 GSCAGCGTGTCTCTCTGACTCCACCTCCGCTGTCTTCCAAAGGCAAGTGCATCAAGGCT 2040
QY 1972 GGCCTGTGATGGAACTCCCGCAGGCGCTCAAGAAGAGATTGCAAGTGTGGGGTGTGGGGGA 2031
Db 2041 GGCCTGTGATGGAACTCCCGCAGGCGCTCAAGAAGAGATTGCAAGTGTGGGGTGTGGGGGA 2100
QY 2032 GACATAGAGCTGCAAGAAGTGAAGTGTCTTCCAAAGGCAAGTGCATGCTACAAT 2091
Db 2101 GACATAGAGCTGCAAGAAGTGAAGTGTCTTCCAAAGGCAAGTGCATGCTACAAT 2160
QY 2092 TTCGTGTGGCCATCCCGCAGGCGCTCAAGCATCGACATCGGCGCAGCGGTTCACAAA 2151
Db 2161 TTCGTGTGGCCATCCCGCAGGCGCTCAAGCATCGACATCGGCGCAGCGGTTCACAAA 2220
QY 2152 GGGCTGTGATGGGATGACAACTACTGCTCTGAAGAACAGCAGCAAGTACCTGCTC 2211
Db 2221 GGGCTGTGATGGGATGACAACTACTGCTCTGAAGAACAGCAGCAAGTACCTGCTC 2280
QY 2212 AACGGGCAATTCGTGTGTGCTGCGGCTGGAGCGGACCTGTGTGTAAGGGCAGTCTGCTG 2271
Db 2281 AACGGGCAATTCGTGTGTGCTGCGGCTGGAGCGGACCTGTGTGTAAGGGCAGTCTGCTG 2340
QY 2272 CGGTACAGCGGACGCGGACAGCGGTGGAGAGCTTCAGGCTTCCCGGCCATCTCGAG 2331
Db 2341 CGGTACAGCGGACGCGGACAGCGGTGGAGAGCTTCAGGCTTCCCGGCCATCTCGAG 2400
QY 2332 CGGCTGACCGTGGAGGCTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCCG 2385
Db 2401 CGGCTGACCGTGGAGGCTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCCG 2454
```

```
RESULT 10
US-10-163-316-1
; Sequence 1, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI01-025PIRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2940)
; OTHER INFORMATION: n = A,T,C or G
US-10-163-316-1
```

```
Query Match 80.5%; Score 2298; DB 41; Length 2940;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Caps 1;
```



```
Db 2632 TTCGTGGTGGCCATCCCGCAGCGGCTCAAGCATCGACATCCGCCAGCGCGTTACAAA 2691
QY 2152 GGCTCATCGGGATGACAACTACCTGGCTCTGAAGACAGCAAGGCAAGTACCTGCTC 2211
Db 2692 GGCTCATCGGGATGACAACTACCTGGCTCTGAAGACAGCAAGGCAAGTACCTGCTC 2751
QY 2212 AACGGGCATTTTCGTGGTGTGCGGCTGGAGCGGGACCTGGTGGTGAAGGCGAGTCTGCTG 2271
Db 2752 AACGGGCATTTTCGTGGTGTGCGGCTGGAGCGGGACCTGGTGGTGAAGGCGAGTCTGCTG 2811
QY 2272 CGGTACAGCGGCACGGGCACAGCGGTGGAGAGCCCTGCAGGCTTCCGGGCCATCCTGGAG 2331
Db 2812 CGGTACAGCGGCACGGGCACAGCGGTGGAGAGCCCTGCAGGCTTCCGGGCCATCCTGGAG 2871
QY 2332 CGGCTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACGCCCGGGTCCGC 2385
Db 2872 CGGCTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACGCCCGGGACCTGC 2925

RESULT 11
US-60-297-863-1
; Sequence 1, Application US/60297863
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses
; FILE OF INVENTION: Therefor
; FILE REFERENCE: MPT2001-025P1(M)
; CURRENT APPLICATION NUMBER: US/60/297,863
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472)...(2941)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2940)
; OTHER INFORMATION: n = A,T,C or G
US-60-297-863-1

Query Match 80.5%; Score 2298; DB 73; Length 2940;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2380; Conservative 0; Mismatches 5; Indels 69; Gaps 1;

QY 1 ATGCTTCTGCTGGGCATCCTAACCTGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG 60
Db 472 ATGCTTCTGCTGGGCATCCTAACCTGCTTTCGCCGGCGCAACCGCTGAGGCTCTGAG 531
QY 61 CCAGAGCGGAGGTAGTGTCCCATCCGACTGGACCGCGGACATTAACGGCCGCGGTAC 120
Db 532 CCAGAGCGGAGGTAGTGTCCCATCCGACTGGACCGCGGACATTAACGGCCGCGGTAC 591
QY 121 TACTGGCGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCATTT 180
Db 592 TACTGGCGGGTCCGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCATTT 651
QY 181 CAGGAGGACTTTTACCTACACTGACCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
Db 652 CAGGAGGACTTTTACCTACACTGACCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 711
QY 241 ACTGAGCATCTGGGCTCCCGCTCCAGGGGTACCGGGGCTCTTCAGACCTTCGACGC 300
Db 712 ACTGAGCATCTGGGCTCCCGCTCCAGGGGTACCGGGGCTCTTCAGACCTTCGACGC 771
QY 301 TGCTTCTATTCTGGGAGGTGAACCGGAGCCGAGCTCGTTCGCTGCTGAGCCTGTC 360
Db 772 TGCTTCTATTCTGGGAGGTGAACCGGAGCCGAGCTCGTTCGCTGCTGAGCCTGTC 831
QY 361 GGGGGTCCCGGAGCTTTGGTACCGAGGCCCGAGTATGTCATTAGCCCGCTGCC 420
Db 420 GGGGGTCCCGGAGCTTTGGTACCGAGGCCCGAGTATGTCATTAGCCCGCTGCC 420
```

```
Db 832 GGGGGTCCCGGAGCTTTGGCTACCGAGGCCCGAGTATGTCATTAGCCCGCTGCC 891
QY 421 AATGCTAGCGCGCGGGGCGGACGCAACAGCGAGGCGCACACCTTTCTCCAGCGCGG 480
Db 892 AATGCTAGCGCGCGGGGCGGACGCAACAGCGAGGCGCACACCTTTCTCCAGCGCGG 951
QY 481 GGTCTTCCGGCGGGGCTTCCGGAGAGCCACCTCTCGCTGCGGGGTGGCTCGGGTGG 540
Db 952 GGTCTTCCGGCGGGGCTTCCGGAGAGCCACCTCTCGCTGCGGGGTGGCTCGGGTGG 1011
QY 541 AACCCGCCCATCTTACGGGCGCTGAGCCCTTACAAGCGCGGGGCGGGCTTCGGGCA 600
Db 1012 AACCCGCCCATCTTACGGGCGCTGAGCCCTTACAAGCGCGGGGCGGGCTTCGGGCA 1071
QY 601 AGTCGTAGCGCGCGAGGTCTGGCGCGGCGCAAGGTTTCGTGTCTATCCGCGGTACGTG 660
Db 1072 AGTCGTAGCGCGCGAGGTCTGGCGCGGCGCAAGGTTTCGTGTCTATCCGCGGTACGTG 1131
QY 661 GAGACGCTGGTGTGCGCGGAGAGTCAATGTCGAAGTTCACAGCGCGGCGGACCTTGGACAT 720
Db 1132 GAGACGCTGGTGTGCGCGGAGAGTCAATGTCGAAGTTCACAGCGCGGCGGACCTTGGACAT 1191
QY 721 TATCTGCTGACGCTGTGCGCAACGCGCGCGGCGGACTCTTACCGCCATCCAGCATCCTCAAC 780
Db 1192 TATCTGCTGACGCTGTGCGCAACGCGCGCGGCGGACTCTTACCGCCATCCAGCATCCTCAAC 1251
QY 781 CCCATCAACATCGTTTGGTCAAGGTCTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 840
Db 1252 CCCATCAACATCGTTTGGTCAAGGTCTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 1311
QY 841 GTCACCGGCAATCGGGCCCTGAGCGCTGCGCAACTTCTGTGCCCTGGCAGAGAAGAGCTGAAC 900
Db 1312 GTCACCGGCAATCGGGCCCTGAGCGCTGCGCAACTTCTGTGCCCTGGCAGAGAAGAGCTGAAC 1371
QY 901 AAAGTGAGTGACAAGCACCCCGAGTACTGGGACACTGCCATCTCTTCCAGGCGAGGAC 960
Db 1372 AAAGTGAGTGACAAGCACCCCGAGTACTGGGACACTGCCATCTCTTCCAGGCGAGGAC 1431
QY 961 CTGTGTGAGCGCACCACTGTGACACCCCTGGGATGGGTGATGTGGTACCATGTGTGAC 1020
Db 1432 CTGTGTGAGCGCACCACTGTGACACCCCTGGGATGGGTGATGTGGTACCATGTGTGAC 1491
QY 1021 CCCAAGAAAGCTGCTCTGTCTATTGAGGAGGATGGGCTTCCATCAGCTTCCACCTGCC 1080
Db 1492 CCCAAGAAAGCTGCTCTGTCTATTGAGGAGGATGGGCTTCCATCAGCTTCCACCTGCC 1551
QY 1081 CAGAGCTGGGCGACAGTGTTCACATGCCCATGCCCATGTGAAAGTGTGTGAGGAGGTG 1140
Db 1552 CAGAGCTGGGCGACAGTGTTCACATGCCCATGCCCATGTGAAAGTGTGTGAGGAGGTG 1611
QY 1141 TTTGGGAAGCTCCGAGCCACCATGATGTCCCGACCTCATCCAGATCGACCTGCC 1200
Db 1612 TTTGGGAAGCTCCGAGCCACCATGATGTCCCGACCTCATCCAGATCGACCTGCC 1671
QY 1201 AACCCCTGTGTAGCTGTGCTGCTGCTATCACCAGCTTCTCTGACAGCGGCGACGCT 1260
Db 1672 AACCCCTGTGTAGCTGTGCTGCTGCTATCACCAGCTTCTCTGACAGCGGCGACGCT 1731
QY 1261 GACTGCTCTGTGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGCCGGCGCC 1320
Db 1732 GACTGCTCTGTGACCAACCCAGAGCCCATCTCCCTGCCCGAGGATCTGCCGGCGCC 1791
QY 1321 AGCTACACCTGAGCGAGCTGCGAGCTGGCTTTTGGGCTGGGCTCCAGCCCTGTCTCT 1380
Db 1792 AGCTACACCTGAGCGAGCTGCGAGCTGGCTTTTGGGCTGGGCTCCAGCCCTGTCTCT 1851
QY 1381 TACATCAGTACTGTACCAAGCTGTGTGCTGCACCGGGAAGCCAAAGGACAGATGTGTGC 1440
Db 1852 TACATCAGTACTGTACCAAGCTGTGTGCTGCACCGGGAAGCCAAAGGACAGATGTGTGC 1911
QY 1441 CAGACCCGCACTTCCCTTGGGCGGATGCCAGCTGTGGCGAGGCAAGCTTGCCTC 1500
Db 1912 CAGACCCGCACTTCCCTTGGGCGGATGCCAGCTGTGGCGAGGCAAGCTTGCCTC 1971
```


Db 842 GTACCGGCAATGCGCCCTGACGTGCGCAACTTCTGTGCTGCGAAGAAGCTGAAC 901
QY 901 AAAGTCAGTGACAAGCACCCCGAGTACTGGGACACTGCCATCTCTTACC----- 951
Db 902 AAAGTCAGTGACAAGCACCCCGAGTACTGGGACACTGCCATCTCTTACCAGGCAAGTT 961
QY 952 ----- 951
Db 962 CCAGGGGAGAGCCCTCTCCACGCTCCGCGGAGCGGCGCTACAGTGCATCTGGCCCAAT 1021
QY 952 ----- 951
Db 1022 GGAGGAGAGCCCTGCGCTTTCCGAAGGTGTGTGGCTTGGGCGGCGCAATCAGCGCCCTCTGG 1081
QY 952 ----- 951
Db 1082 ATCAGGCGCGAGGGCGGAACCCAGGAAGTTGCCGCGCTGGAGCTGCAATTTGTGTGCC 1141
QY 952 ----- 951
Db 1142 AAGACCGATAGGAGACGCGGTGAGGATGGTGTGGAGAGCGGGAAGCGCCACCCCTA 1201
QY 952 ----- 951
Db 1202 TTGTATGGCGGCTGAGTCTTCTCGGACACCTCTGAGGTCTGCTTTCAAGGTTGAGA 1261
QY 952 ----- 951
Db 1262 ACTGAAGGTGATCCAAAGGTGGCTCCCAATCCTAAATTTATTTATAGAAGAGGACAGTTA 1321
QY 952 ----- 951
Db 1322 GAAATAACAGTCCAGGAGGATTAGGTTCCTGGCCAGAGATGATTCAGGATCATCTGGA 1381
QY 952 ----- 951
Db 1382 AAAGTGGCTGCAGCTTACAAAGAAAACAGCCAAATTTGCTGGAAATCGTGAGAGCAGGACTG 1441
QY 964 TGTGAGGACACACCTGTGACACCTGGCATGTGCTGATGTGGGTACCATCTGTGACCCC 1023
Db 1442 TGTGAGGACACACCTGTGACACCTGGCATGTGCTGATGTGGGTACCATCTGTGACCCC 1501
QY 1024 AAGAAAGCTGTGTGTCATTGAGAGAGATGGGCTTCCATCAGCCTTCACTACCTGCCAC 1083
Db 1502 AAGAAAGCTGTGTGTCATTGAGAGAGATGGGCTTCCATCAGCCTTCACTACCTGCCAC 1561
QY 1084 GAGCTGGGCCACGTGTCAACATGCCCATGCCATGACAATGTGAAGTCTGTGAGGAGGTGTT 1143
Db 1562 GAGCTGGGCCACGTGTCAACATGCCCATGCCATGACAATGTGAAGTCTGTGAGGAGGTGTT 1621
QY 1144 GGAAGCTCCGAGCCAAACCATGATGTCCCGGACCTCATCCAGATCGACCGTGGCAAC 1203
Db 1622 GGAAGCTCCGAGCCAAACCATGATGTCCCGGACCTCATCCAGATCGACCGTGGCAAC 1681
QY 1204 CCCTGGTCAGCTGAGTGTGCCATCATCACCAGCTTCCTGGACAGCGGCGAGGTGAC 1263
Db 1682 CCCTGGTCAGCTGAGTGTGCCATCATCACCAGCTTCCTGGACAGCGGCGAGGTGAC 1741
QY 1264 TGCCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCGAGGATCTGCCGGGCGGACG 1323
Db 1742 TGCCTCTCTGGACCAACCCAGCAAGCCCATCTCCCTGCCGAGGATCTGCCGGGCGGACG 1801
QY 1324 TACACCTTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTGCTTAC 1383
Db 1802 TACACCTTGAGCCAGCAGTGGAGCTGGCTTTTGGCGTGGGCTCCAAAGCCCTGTGCTTAC 1861
QY 1384 ATGACGTAAGTCCACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGGTGTGCAG 1443
Db 1862 ATGACGTAAGTCCACCAAGCTGTGTGTCACCGGGAAGGCCAAGGACAGATGGTGTGCAG 1921
QY 1444 ACCGCGCACTTCCCTGGCGGATGGCACCAGCTGTGGCGAGGCGCAAGCTCTGCCTCAA 1503

Db 1922 ACCGCGCACTTCCCTGGCGGATGGCACCAGCTGTGGCGAGGCAAGCTCTGCCTCAA 1981
QY 1504 GGGGCTTGGCTGGAGAGACACAACCTCAACAAGCACAG----- 1541
Db 1982 GGGGCTTGGCTGGAGAGACACAACCTCAACAAGCACAGGGCTTTGACTGACATCATTTCT 2041
QY 1542 -----GGTGGATGGTTC 1554
Db 2042 CCAAAACAACCTCTATTAAAGGCTACCAAAATGGCGCTGCACACTAGAAAGGTGGATGGTTC 2101
QY 1555 TGGCCAAATGGGATCCCTATGGCCCTCTGTCGCGCACATGTGTGGGGGCGCTGCACACTG 1614
Db 2102 TGGCCAAATGGGATCCCTATGGCCCTCTGTCGCGCACATGTGTGGGGGCGCTGCAGCTG 2161
QY 1615 GCCAGGAGGAGTGCACCAACCCCTGCCACCGGCGCAAGTACTCGGAGGGAGTGG 1674
Db 2162 GCCAGGAGGAGTGCACCAACCCCTGCCAACCGGCGCAAGTACTCGGAGGGAGTGG 2221
QY 1675 AGGTGAAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAAGAGC 1734
Db 2222 AGGTGAAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAAGAGC 2281
QY 1735 TTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACACAGCACCAACCGGCTCCT 1794
Db 2282 TTCGGGAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACACAGCACCAACCGGCTCCT 2341
QY 1795 CTGCGCTGTCATGGGTGCCAAGTACTCCGGGCTGTCTCCCGGGGACAAAGTGCACACTC 1854
Db 2342 CTGCGCTGTCATGGGTGCCAAGTACTCCGGGCTGTCTCCCGGGGACAAAGTGCACACTC 2401
QY 1855 ATCTGCGGAGCCCAATGCACTGCTTCTATGTGCTGCAACCCAGGTGGTGGACGGC 1914
Db 2402 ATCTGCGGAGCCCAATGCACTGCTTCTATGTGCTGCAACCCAGGTGGTGGACGGC 2461
QY 1915 ACCTGTGCTCTCCTGACTCCACTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGC 1974
Db 2462 ACCTGTGCTCTCCTGACTCCACTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGC 2521
QY 1975 TGTGATGGGAACCTGGGCTCCAGAGAGATTCGACAAGTGTGGGTGTGTGGGGGAGAC 2034
Db 2522 TGTGATGGGAACCTGGGCTCCAGAGAGATTCGACAAGTGTGGGTGTGTGGGGGAGAC 2581
QY 2035 AATAAGAGCTGCAAGAAGTGTGACTGCTTCCACCAAGCCCATGCTGCTGCTACAAATTC 2094
Db 2582 AATAAGAGCTGCAAGAAGTGTGACTGCTTCCACCAAGCCCATGCTGCTGCTACAAATTC 2641
QY 2095 GTGGTGGCCATCCCGGAGCGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGG 2154
Db 2642 GTGGTGGCCATCCCGGAGCGCTCAAGCATCGACATCCGCGAGCGGTTTACAAAGG 2701
QY 2155 CTGATCGGGATGACAACTACCTGGCTCTGAAGAAGCAAGGCAAGTACCTGCTCAA 2214
Db 2702 CTGATCGGGATGACAACTACCTGGCTCTGAAGAAGCAAGGCAAGTACCTGCTCAA 2761
QY 2215 GGGCATTTGCTGTGCTGGCGGTGGAGCGGACCTGCTGGTGAAGGCAAGTCTGTGCGG 2274
Db 2762 GGGCATTTGCTGTGCTGGCGGTGGAGCGGACCTGCTGGTGAAGGCAAGTCTGTGCGG 2821
QY 2275 TACAGCGGACCGGACACAGCGGTGGAGAGCTGTGAGGCTTCCCGGCCCATCTGTGAGCGG 2334
Db 2822 TACAGCGGACCGGACACAGCGGTGGAGAGCTGTGAGGCTTCCCGGCCCATCTGTGAGCGG 2881
QY 2335 CTGACGCTGGAGTCTCTCCGTGGGGAAGATGACACCGCGGCTGCGGCTTCCCGGCCCATCTGTGAGCGG 2394
Db 2882 CTGACGCTGGAGTCTCTCCGTGGGGAAGATGACACCGCGGCTGCGGCTTCCCGGCCCATCTGTGAGCGG 2941
QY 2395 TATCTGCCCAAAGAGCTCGGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCC 2454
Db 2942 TATCTGCCCAAAGAGCTCGGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCC 3001
QY 2455 TCTGTCTTGCACACAGCGCTCTCAGCTCTCCAAACAGGTGGAGAGCGGACGACAGG 2514
Db 3002 TCTGTCTTGCACACAGCGCTCTCAGCTCTCCAAACAGGTGGAGAGCGGACGACAGG 3061

QY 973 ACCACCTGTGACACCCCTGGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAGC 1032
Db 1681 ACCACCTGTGACACCCCTGGGCATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAAGC 1740
QY 1033 TGCTCTGTGATGAGGACGATGGGCTTCATCAGCCCTTACACACTGCCCCAAGAGTGGG 1092
Db 1741 TGCTCTGTGATGAGGACGATGGGCTTCATCAGCCCTTACACACTGCCCCAAGAGTGGG 1800
QY 1093 CACGTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAGGAGGTGTTTGGGAAGCTC 1152
Db 1801 CACGTGTTCAACATGCCCATGACAAATGTGAAAGTCTGTGAGGAGGTGTTTGGGAAGCTC 1860
QY 1153 CGAGCAACCAACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGCAACCCCTGTGTCA 1212
Db 1861 CGAGCAACCAACATGATGTCCCGACCCCTCATCCAGATCGACCGTGGCAACCCCTGTGTCA 1920
QY 1213 GCCTGCACTGTGATCATCACCAGCTTCTTGACAGCGGGGACGGTGAAGTGGCTCCTG 1272
Db 1921 GCCTGCACTGTGATCATCACCAGCTTCTTGACAGCGGGGACGGTGAAGTGGCTCCTG 1980
QY 1273 GACCAACCAAGCAAGCCCATCTCCCTGCCGAGGATCTGCCGGGCGCCAGCTACACCCCTG 1332
Db 1981 GACCAACCAAGCAAGCCCATCTCCCTGCCGAGGATCTGCCGGGCGCCAGCTACACCCCTG 2040
QY 1333 AGCCAGAGTGGGAGCTGGCTTTTGGCGTGGGCTTCCAGCCCTGTCTTACATGACATAC 1392
Db 2041 AGCCAGAGTGGGAGCTGGCTTTTGGCGTGGGCTTCCAGCCCTGTCTTACATGACATAC 2100
QY 1393 TGCACCAAGCTGTGTGCACCGAAGCCCAAGGACAGATGGTGTGCCAGACCCGCCAC 1452
Db 2101 TGCACCAAGCTGTGTGCACCGAAGCCCAAGGACAGATGGTGTGCCAGACCCGCCAC 2160
QY 1453 TTCCCTGGGCGGATGGCACCAGCTGTGGGAGGGCAAGCTCTGCCCTCAAGGGGCGCTGC 1512
Db 2161 TTCCCTGGGCGGATGGCACCAGCTGTGGGAGGGCAAGCTCTGCCCTCAAGGGGCGCTGC 2220
QY 1513 GTGAGAGACACACCTCAACACACAG----- 1541
Db 2221 GTGAGAGACACACCTCAACACACAGGCTTTCACTGATGACCTTTCTCTCATTTCC 2280
QY 1542 ----- 1541
Db 2281 CGACGGGGGAAAAACCAACAGGTTCCGGAGGAGTTGTACAGCTTGTCTTCCAAAT 2340
QY 1542 ----- 1541
Db 2341 TCCATCTTCCCACTTTTCTCTGCTGCCCATTTCCAAACAGGCTTTGACTGACATATTCT 2400
QY 1542 -----GGTGGATGGTTCC 1554
Db 2401 CCAAAACAACTCTATTAAAGCTACCAAAATGGGCTGCACACTACGAAGGTTGATGGTTCC 2460
QY 1555 TGGCCCAATGGGATCCCTATGGCCCTGCTCGCGACATGTGTGGGGGCGGTGACAGCTG 1614
Db 2461 TGGCCCAATGGGATCCCTATGGCCCTGCTCGCGACATGTGTGGGGGCGGTGACAGCTG 2520
QY 1615 GCGAGGAGGAGTGCACCAACCCCAACCCCTGCGCAAGGGGGCAAGTACTGAGGGAGTG 1674
Db 2521 GCGAGGAGGAGTGCACCAACCCCAACCCCTGCGCAAGGGGGCAAGTACTGAGGGAGTG 2580
QY 1675 AGGTGAAATACCGATCTGCAATCTGAGCCCTGAGCCCTGCCCCAGTTCAGCCCTCGGAAGAGC 1734
Db 2581 AGGTGAAATACCGATCTGCAATCTGAGCCCTGAGCCCTGCCCCAGTTCAGCCCTCGGAAGAGC 2840
QY 1735 TTCGGGAGGAGTGTGAGGCTTTTCAAGCGCTTCAACACAGCAGCACCACCTGGCTCACT 1794
Db 2641 TTCGGGAGGAGTGTGAGGCTTTTCAAGCGCTTCAACACAGCAGCACCACCTGGCTCACT 2700
QY 1795 CTGCGCTGGATGGGTGCCAAGTACTCGGCGTGTCTCCCGGGACAAAGTGCAGAGCTC 1854
Db 2701 CTGCGCTGGATGGGTGCCAAGTACTCGGCGTGTCTCCCGGGACAAAGTGCAGAGCTC 2760
QY 1855 ATCTGCCGAGCAATGCACTGCTACTTCTATGTCTGCGCACCACAAAGGTTGGTGGACGGC 1914

Db 2761 ATCTGCCGAGCCATGCACTGGCTACTTCTATGTGTGCAACCCCAAGGTTGGTGAAGC 2820
QY 1915 ACCTGTGCTCTCTCTGACTCCACTCCGCTGTGTCTCAAGGCAAGTGCATCAAGGCTGGC 1974
Db 2821 ACCTGTGCTCTCTCTGACTCCACTCCGCTGTGTCTCAAGGCAAGTGCATCAAGGCTGGC 2880
QY 1975 TGTGATGGGAACCTGGGCTCCAAAGAGAGATTGCAAGTGTGGGGTGTGTGGGGGAGAC 2034
Db 2881 TGTGATGGGAACCTGGGCTCCAAAGAGAGATTGCAAGTGTGGGGTGTGTGGGGGAGAC 2940
QY 2035 AATAAGAGCTGCAAGAGGTTGACTGGACTCTTCCACCAAGCCCATGATGGCTACAAATTC 2094
Db 2941 AATAAGAGCTGCAAGAGGTTGACTGGACTCTTCCACCAAGCCCATGATGGCTACAAATTC 3000
QY 2095 GTGGTGGCCATCCCGCAGCGCTCAAGCATCGACATCGCCAGCGCGGTTACAAAGG 2154
Db 3001 GTGGTGGCCATCCCGCAGCGCTCAAGCATCGACATCGCCAGCGCGGTTACAAAGG 3060
QY 2155 CTGATCGGGATCACAACTACCTGGCTCTGGAAGAACAGCAAGGCAAGTACCTGTAAAC 2214
Db 3061 CTGATCGGGATCACAACTACCTGGCTCTGGAAGAACAGCAAGGCAAGTACCTGTAAAC 3120
QY 2215 GGGCATTTCTGTGTGTGGCGTGGAGCGGACTCTGTGTGAAGGCAAGTCTGTGCGG 2274
Db 3121 GGGCATTTCTGTGTGTGGCGTGGAGCGGACTCTGTGTGAAGGCAAGTCTGTGCGG 3180
QY 2275 TACAGCGGCACGGGCACAGCGTGGAGAGCTTCCAGGCTTCCCGGCCCATCTTGAAGCG 2334
Db 3181 TACAGCGGCACGGGCACAGCGTGGAGAGCTTCCAGGCTTCCCGGCCCATCTTGAAGCG 3240
QY 2335 CTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCGCTACTCTTTC 2394
Db 3241 CTGACCGTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCGCTACTCTTTC 3300
QY 2395 TATCTGCCAAAGAGCTCGGGAGGACAAAGTCTCTCTATCCCAAGGACCCCGGGAGCC 2454
Db 3301 TATCTGCCAAAGAGCTCGGGAGGACAAAGTCTCTCTATCCCAAGGACCCCGGGAGCC 3360
QY 2455 TCTGTCTTGCACAAACAGCTCTCTCAAGCTTCCAAACAGGTGGAGAGCGGACACAGG 2514
Db 3361 TCTGTCTTGCACAAACAGCTCTCTCAAGCTTCCAAACAGGTGGAGAGCGGACACAGG 3420
QY 2515 CCCCCGTGACGCTGGGTGGCTGGAGCTGGGGCGTGTCTCCGAGCTGCGGAGTGGCGAGTGGC 2574
Db 3421 CCCCCGTGACGCTGGGTGGCTGGAGCTGGGGCGTGTCTCCGAGCTGCGGAGTGGCGAGTGGC 3480
QY 2575 CTGCAAGAGCGGGCGGTGGACTTCTCGGGCTCCCGGGGACGCGACGCTCCCTGCT 2634
Db 3481 CTGCAAGAGCGGGCGGTGGACTTCCCGGGCTCCCGGGGACGCGACGCTCCCTGCT 3540
QY 2635 GATGAGCCCATCGGCGCTGGAGACACAAAGCTTGGGGGAGCCCTGCCACCTGAGAG 2694
Db 3541 GATGAGCCCATCGGCGCTGGAGACACAAAGCTTGGGGGAGCCCTGCCACCTGAGAG 3600
QY 2695 CTCAGCGCTGTCTACCTCTCTCAAGAGCTGCGCGCGGGGATTTCAGAGGCGCTCACTC 2754
Db 3601 CTCAGCGCTGTCTACCTCTCTCAAGAGCTGCGCGCGGGGATTTCAGAGGCGCTCACTC 3660
QY 2755 AAGTGTGGGCGCAGGAGCGGCTGTGGCCCCGGGACCAAGTGCACCTTGCACCGAAG 2814
Db 3661 AAGTGTGGGCGCAGGAGCGGCTGTGGCCCCGGGACCAAGTGCACCTTGCACCGAAG 3720
QY 2815 CCCCCAGGAGTGGACTTCTGGCTCTGTAGGCGGCTCTGA 2853
Db 3721 CCCCCAGGAGTGGACTTCTGGCTCTGTAGGCGGCTCTGA 3759

RESULT 14

US-60-230-435-2152

; Sequence 2152, Application US/60230435

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2152
; LENGTH: 3625
; TYPE: DNA
; ORGANISM: HUMAN
; US-60-230-435-2152

Query Match 66.9%; Score 1907.8; DB 67; Length 3625;
Best Local Similarity 77.0%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 237; Indels 546; Gaps 2;

Qy	1	ATGCTTCTGCTGGCATCTTAACCTTGGCTTTCCGCGGCGGAACCCGCTGGAGGCTCTGAG	60
Db	177	ATGCTTCTGCTGGCATCTTAACCTTGGCTTTCCGCGGCGGAACCCGCTGGAGGCTCTGAG	236
Qy	61	CCAGAGCGGAGGTAGTCTGTTCCATCCGACTGGACCCCGGACATTAACGGCCCGCTAC	120
Db	237	CCAGAGCGGAGGTAGTCTGTTCCATCCGACTGGACCCCGGACATTAACGGCCCGCTAC	296
Qy	121	TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT	180
Db	297	TACTGGCGGGTCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT	356
Qy	181	CAGGAGACTTTTACCTTACACTGACGCGGATGCTGAGTCTTGGCTCCCGCTTCTCC	240
Db	357	CAGGAGACTTTTACCTTACACTGACGCGGATGCTGAGTCTTGGCTCCCGCTTCTCC	416
Qy	241	ACTGAGACTTGGCGTCCCGCTCCAGGGCTCACCGGGGCTTCAGACCTGCGACGC	300
Db	417	ACTGAGACTTGGCGTCCCGCTCCAGGGCTCACCGGGGCTTCAGACCTGCGACGC	476
Qy	301	TGCTTTCTATTCTGGGAGCTGAACCGCGAGCCGAGTCTGCTGCTGCTGAGCTGTGC	360
Db	477	TGCTTTCTATTCTGGGAGCTGAACCGCGAGCCGAGTCTGCTGCTGCTGAGCTGTGC	536
Qy	361	GGGGGCTCCGCGAGCTTTGGCTACCGAGCGCCGAGTATGTCATTAGCCGCTGCC	420
Db	537	GGGGGCTCCGCGAGCTTTGGCTACCGAGCGCCGAGTATGTCATTAGCCGCTGCC	596
Qy	421	AATCTAGCGCGCGCGGCGAGCAACAGCAGCGGCGACAGCTTCTCCAGCGCGG	480
Db	597	AATCTAGCGCGCGCGGCGAGCAACAGCAGCGGCGACAGCTTCTCCAGCGCGG	656
Qy	481	GGTGTTCGGCGGGCTTCGGAGACCCACCTCTCGCTCGCGGGTGGCTCGGGCTGG	540
Db	657	GGTGTTCGGCGGGCTTCGGAGACCCACCTCTCGCTCGCGGGTGGCTCGGGCTGG	716
Qy	541	AACCCGCGCATCTACGGGCGCTGGACCTTACAGCCGCGGGGCGGCTTCGGGGAG	600
Db	717	AACCCGCGCATCTACGGGCGCTGGACCTTACAGCCGCGGGGCGGCTTCGGGGAG	776
Qy	601	AGTGTAGCGCGCGAGTCTGGCGCGCCAAAGGTTTCTGTTCTATCCCGCGGTACGTG	660
Db	777	AGTGTAGCGCGCGAGTCTGGCGCGCCAAAGGTTTCTGTTCTATCCCGCGGTACGTG	836
Qy	661	GAGAGCTGTGGTCCGGAGAGTCAATGGTCAAGTTTCCAGCGCGCGGACCTGGAAACAT	720
Db	837	GAGAGCTGTGGTCCGGAGAGTCAATGGTCAAGTTTCCAGCGCGCGGACCTGGAAACAT	896
Qy	721	TATCTGTAGCTGTGGTCCGGAGAGTCAATGGTCAAGTTTCCAGCGCGCGGACCTGGAAACAT	780
Db	897	TATCTGTAGCTGTGGTCCGGAGAGTCAATGGTCAAGTTTCCAGCGCGCGGACCTGGAAACAT	956
Qy	781	CCATCAACATCGTGTGGTCAAGTGTGCTTCTTAGAGATCGTACTCCGGGCGCAAG	840
Db	957	CCATCAACATCGTGTGGTCAAGTGTGCTTCTTAGAGATCGTACTCCGGGCGCAAG	1016

; Qy 841 GTCACCGGCAATGGGCGCTGACGCTCGGCAACTTCTGTGCTGGCGAAGAGAGCTGAAC 900
; Db 1017 GTCACCGGCAATGGGCGCTGACGCTCGGCAACTTCTGTGCTGGCGAAGAGAGCTGAAC 1076
; Qy 901 AAGTGTAGTACAGACACCCGAGTACTGGGACACTGCTCTTTCACAGGACGAC 960
; Db 1077 AAGTGTAGTACAGACACCCGAGTACTGGGACACTGCTCTTTCACAGGACGAGTG 1136
; Qy 961 CTGTGTGAGGACACACCTGTGACACCCCTGGGACATGCTGATGTGGG----- 1007
; Db 1137 AGTTGATCGCGTCACTTTGACCCAGATAGTCCCGTCTTTAGGGCATCTGGGCCAATT 1196
; Qy 1008 ----- 1007
; Db 1197 GGAGGAGAGCTGCGCTTTCCGAAGTGTGGCTGGCGGCGCAATCAGCGCTCC1GG 1256
; Qy 1008 ----- 1007
; Db 1257 ATCAGGCGCCGAGGGCGCGGAACCCAGGAAGTTCGCCGCCCGGAGCTGCAGTTTGTCTCC 1316
; Qy 1008 ----- 1007
; Db 1317 AAGACCGATAGGAGACGCGGTGAGATGGTGTGGAGAGGCGGGAACGGCCACCCCTA 1376
; Qy 1008 ----- 1007
; Db 1377 TTGTATGGCGGCTGAGTCTTTCGGACACCTCCTGAGGTCCTTCAAGGGTGTAGA 1436
; Qy 1008 ----- 1007
; Db 1437 ACTGAAGTGTATCAAGGTCAAGCTGAGCTTCTCTCGGGTAAACAGCTTGTCCCT 1496
; Qy 1008 ----- 1007
; Db 1497 CTCGTAAACAGTCCAGGAGGATTAAAGTTTCTGCGCAAGATGATTCAGGATCATCTGA 1556
; Qy 1008 ----- 1007
; Db 1557 AAAGTGGCTGCAGCTTACAAAGAAAGCAATGCTGGGAATGCTGAGGACGAGTAGGG 1616
; Qy 1008 -----TACCATGTGTGACCCC 1023
; Db 1617 GCTGGTTAGGATACCTTACACCTTCTGGAGGAGAGGACAGATACCATGTGTGACCCC 1676
; Qy 1024 AAGAGAAGTGTCTGTATGAGGACGATGGGCTTCCATCAGCCTTCCACACTGCCCCAC 1083
; Db 1677 AAGAGAAGTGTCTGTATGAGGACGATGGGCTTCCATCAGCCTTCCACACTGCCCCAC 1736
; Qy 1084 GAGCTGGGCGGCTGTTCAACATGCCCATGACATGTGAAAGTCTGTGAGGAGGTGTT 1143
; Db 1737 GAGCTGGGTAAGGCTGGATAGCTCCTCTGGGGTCTTCTGGGTTTGCTGGGAGCTTT 1796
; Qy 1144 GGAAGCTCCGAGCAACACATGATGCCCGACCCCTCATCCAGATCGACCGTGCACAC 1203
; Db 1797 GGAAGCTCCGAGCAACACATGATGCCCGACCCCTCATCCAGATCGACCGTGCACAC 1856
; Qy 1204 CCTGTGAGCTGAGTGTCTCCATCATCAGGACTTCTGGACAGGCGGACGAGTGCAC 1263
; Db 1857 CCTGTGAGCTGAGTGTCTCCATCATCAGGACTTCTGGACAGGCGGACGAGTGCAG 1916
; Qy 1264 TGCTCTCGACCAACCCAGCAAGCCCATCTCCCTGCCGAGGATCTGCCGGGCGCCAGC 1323
; Db 1917 CCAGAGCGGGAGGGCAATGAGCGCGCTCGAGGGGGCTTTGCTGCCGGGCGCCAGC 1976
; Qy 1324 TACACCTGAGCAGCAGTGCAGCTGAGTGGCTTTGGCGTGGGCTCCAAGCCCTGTCTTAC 1383
; Db 1977 TACACCTGAGCAGCAGTGCAGCTGAGTGGCTTTGGCGTGGGCTCCAAGCCCTGTCTTAC 2036
; Qy 1384 ATGCAGTACTGCACCAAGCTGTGTGACCGGGAAGGCAAGGACAGATGTTGTGCCAG 1443
; Db 2037 ATGCAGTACTGCACCAAGCTGTGTGACCGGGAAGGCAAGGACAGATGTTGTGCCAG 2096

QY 1444 ACCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGGCGAGGCAAGCTCTGCTCAAA 1503
DB 2097 ACCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGGCGAGGCAAGCTCTGCTCAAA 2156
QY 1504 GGGGCTCGGTGGAGAGACACAACTCAACAGACACAGGCTGGATGCTTCTGGGCC--- 1560
DB 2157 GGGGCTCGGTGGAGAGACACAACTCAACAGACACAGGCTGGATGCTTCTGGAGTGC 2216
QY 1561 ----- 1560
DB 2217 GCTGGGAGCTGCTGGAGGAGGATGGATGGCTGCACACTAGCAAGGTCAGTGACTGT 2276
QY 1561 -----AAATGGGATCCCTATGTCGCGCAGCTGTGGTGGGGCTGGAGCTG 1614
DB 2277 TACCTTTCTCCATCGTCCAGCGGCTAGCAGCATTTATGTGGTGGGGCGTGCAGCTG 2336
QY 1615 GCAGAGGAGCTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTG 1674
DB 2337 GCCAGAGGAGCTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGAGGGAGTG 2396
QY 1675 AGGTGAATACGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCCTCGGAAAGAGC 1734
DB 2397 AGGTGAATACGATCTGCAATCTGAGCCCTGCGCCAGCTCAGCCTCGGAGGAGAG 2456
QY 1735 TTCCGGGAGGAGCTGTGAGCTTTCAACGGCTTCAACAGCTACACACAGCAACCGGCTCACT 1794
DB 2457 GCAGTGGTGGCTGGCCAGGGAGGTGAGCTGGAGTGAACACACACACCGGCTCACT 2516
QY 1795 CTCGGCTGGCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTCAAGCTC 1854
DB 2517 CTCGGCTGGCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTCAAGCTC 2576
QY 1855 ATCTCCGAGCAANTGGCACTGGCTTCTATGTGCTGCGACCAAGCTGGTGGAGCGC 1914
DB 2577 ATCTCCGAGCAANTGGCACTGGCTTCTATGTGCTGCGACCAAGCTGGTGGAGCGC 2636
QY 1915 ACCTGTGTCTCTGACTCTCCACCTCCCTGTGTGTCTCAAGCAAGTCAAGCTGGC 1974
DB 2637 GGGGCTTGAGAACAAAGTAGGAGCAGGCTTCCGGGGAGCAAGTCAAGCTGGC 2696
QY 1975 TGTGATGGCACTGGCTCCAGAGAGATTCGACAGTGTGGGTGTGGGGAGAGC 2034
DB 2697 TGTGATGGCACTGGCTCCAGAGAGATTCGACAGTGTGGGTGTGGGGAGAGC 2756
QY 2035 AATAAGAGCTGAAGAAGTGACTGGACTTTCACCAAGCCCTGCATGGCTACAATTC 2094
DB 2757 AATAAGAGCTGAAGAAGTGACTGGACTTTCACCAAGCCCTGCATGGCTACAATTC 2816
QY 2095 GTGGTGGCCATCCCGCAGCGCCCTCAAGCATCGACATCGCCAGCGGGTTACAAAGG 2154
DB 2817 GAAGTCTCTGCCAGGGGCAAGAGGAGGAGGTGACATCCCGCAGCGGGTTACAAAGG 2876
QY 2155 CTGATCGGGATGACAACTTACCTGGCTGTGAAGAACAGCAAGCAAGTCACTCTCAAC 2214
DB 2877 CTGATCGGGATGACAACTTACCTGGCTGTGAAGAACAGCAAGCAAGTCACTCTCAAC 2936
QY 2215 GGCATTTCTGTGTGCGCGGTGAGAGCGGACCTGTGTGTGAAGGGCAGTCTGCTGGG 2274
DB 2937 GGCATTTCTGTGTGCGCGGTGAGAGCGGACCTGTGTGTGAAGGGCAGTCTGCTGGG 2996
QY 2275 TACAGCGGACGGGACAGCGGTGGAGAGCCTTGCAGGCTTCCCGGCCATCTCTGGAGCG 2334
DB 2997 TACAGCGGACGGGACAGCGGTGGAGAGCCTTGCAGGCTTCCCGGCCATCTCTGGAGCG 3056
QY 2335 CTGACCGTGGAGTCTCTCTCGGTGGGAAAGATGACACCGCCCGGGTCCGCTACTCTTC 2394
DB 3057 CTGACCGTGGAGTCTCTCTCGGTGGGAAAGATGACACCGCCCGGGTCCGCTACTCTTC 3116
QY 2395 TATCTCCCAAGAGCTCGGAGGACAGTCTCTCTCAACAGGACCCCGGGAGCC 2454
DB 3117 TATCTCCCAAGAGCTCGGAGGACAGTCTCTCTCAACAGGACCCCGGGAGCC 3176
QY 2455 TCTGTCTTCACAAAGCGTCTCTCAGCCCTCTCCACAGGCTGGAGCGCGGAGCAGG 2514

DB 3177 TCTGTCTTCACAAAGCGTCTCTCAGCCTCTCCAAACAGGTGGAGCGCGGACGACAGG 3236
QY 2515 CCCCTGCACGCTGGTGGCTGGCAGCTGGGGCCGCTGCTCCGAGCTCGGAGTGGC 2574
DB 3237 CCCCTGCACGCTGGTGGCTGGCAGCTGGGGCCGCTGCTCCGAGCTCGGAGTGGC 3296
QY 2575 CTGCAAGAGCGGGGCTGGACTGTGGGGCTCCCGGGGAGCGCAGCTCCCTGCTGT 2634
DB 3297 CTGCAAGAGCGGGGCTGGACTGTGGGGCTCCCGGGGAGCGCAGCTCCCTGCTGT 3356
QY 2635 GATGACAGCCCATCGSCCGCTGGAGACACAAAGCCTCGGGGAGCCTGCCCACTGGGAG 2694
DB 3357 GATGACAGCCCATCGSCCGCTGGAGACACAAAGCCTCGGGGAGCCTGCCCACTGGGAG 3416
QY 2695 CTCAGCGCTGTGCTACCTGCTCCAAAGAGCTGCGGCGGGGATTCAGAGCGCTCACT 2754
DB 3417 CTCAGCGCTGTGCTACCTGCTCCAAAGAGCTGCGGCGGGGATTCAGAGCGCTCACT 3476
QY 2755 AAGTGTGGGCGCAGGAGCGGCTGCTGCGCTCTGAGGCCGTGCTGA 2853
DB 3477 AAGTGTGGGCGCAGGAGCGGCTGCTGCGCTCTGAGGCCGTGCTGA 3575
QY 2815 CCCCAGGAGCTGGACTTCTGCTCTGAGGCCGTGCTGA 2853
DB 3537 CCCCAGGAGCTGGACTTCTGCTCTGAGGCCGTGCTGA 3575

RESULT 15

US-10-093-463-29

Sequence 29, Application US/10093463

GENERAL INFORMATION:

APPLICANT: Padigaru, Muralidhara
APPLICANT: Shenoy, Suresh
APPLICANT: Kekuda, Ramesh
APPLICANT: Gusev, Vladimir
APPLICANT: Pochart, Pascal
APPLICANT: Zhong, Mei
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter
APPLICANT: Smithson, Glenda
APPLICANT: Guo, Xiaojia
APPLICANT: Gerlach, Valerie
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Zerhusen, Bryan
APPLICANT: Tchernev, Velizar
APPLICANT: Gangolli, Esha
APPLICANT: Vernet, Corine
APPLICANT: Pena, Carol
APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Patturajan, Meera
APPLICANT: Miller, Charles
APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic
FILE REFERENCE: 21402-290A (Cura 590A)
CURRENT APPLICATION NUMBER: US/10/093,463
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08


```
Db 981 -----GGTGGATGTTCTCTGGCC 999
QY 1561 AAATGGATCCCTATGGCCCTGTCGCGGCACATGTGTGGGGCGCTGCAGCTGGCCAGG 1620
Db 1000 AAATGGATCCCTATGGCCCTGTCGCGGCACATGTGTGGGGCGCTGCAGCTGGCCAGG 1059
QY 1621 AGGCACTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGGAGGAGTGAAGGTG 1680
Db 1060 AGGCACTGCACCAACCCACCCCTGCCAACGGGGCAAGTACTCGGAGGAGTGAAGGTG 1119
QY 1681 AAATACCGATCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCCGAAAGAGTTCGCG 1740
Db 1120 AAATACCGATCTGCAATCTGAGAGCCCTGCCCCAGCTCAGCTCCCGAAAGAGTTCGCG 1179
QY 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTTACAAACAGCAGCACCAACCGGCTCACTCTCGCC 1800
Db 1180 GAGGAGCAGTGTAGGCTTTCAACGGCTTACAAACAGCAGCACCAACCGGCTCACTCTCGCC 1239
QY 1801 GTGGCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTGCAGCTCATCTGC 1860
Db 1240 GTGGCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGGACAAAGTGCAGCTCATCTGC 1299
QY 1861 CGAGCCAATGGCACTGGCTACTTCTATGTGTGGCACCACCAAGGTGTGGAGCGCACGCTG 1920
Db 1300 CGAGCCAATGGCACTGGCTACTTCTATGTGTGGCACCACCAAGGTGTGGAGCGCACGCTG 1359
QY 1921 TGCTCTCTCTACTCCACCTCCCTGTGTGTCCAAAGCCATGCATGGCTACAATTTCTGGTG 1980
Db 1360 TGCTCTCTCTACTCCACCTCCCTGTGTGTCCAAAGCCATGCATGGCTACAATTTCTGGTG 1419
QY 1981 GGGAACTGGCTCCAAAGAGAGATTCGACAGTGTGGGTGTGGGGAGAGCAATAAG 2040
Db 1420 GGGAACTGGCTCCAAAGAGAGATTCGACAGTGTGGGTGTGGGGAGAGCAATAAG 1479
QY 2041 AGTGTCAAGAAGTGACTGGACTCTTACCAAGCCCATGCATGGCTACAATTTCTGGTG 2100
Db 1480 AGTGTCAAGAAGTGACTGGACTCTTACCAAGCCCATGCATGGCTACAATTTCTGGTG 1539
QY 2101 GCCATCCCGCAGCGCCCTCAGCATCGACATCCCGCAGCGGTTACAAAGGCTGATC 2160
Db 1540 GCCATCCCGCAGCGCCCTCAGCATCGACATCCCGCAGCGGTTACAAAGGCTGATC 1599
QY 2161 GGGGATGACAACCTACCTGGCTCTGAAGAACGCCAAGCAAGTACTGCTCAAGGGGCAT 2220
Db 1600 GGGGATGACAACCTACCTGGCTCTGAAGAACGCCAAGCAAGTACTGCTCAAGGGGCAT 1659
QY 2221 TTCGTGGTGTGGCGGTGAGCGGACCTGTGGTGAAGGGCAGTCTGCTCGGGTACAGC 2280
Db 1660 TTCGTGGTGTGGCGGTGAGCGGACCTGTGGTGAAGGGCAGTCTGCTCGGGTACAGC 1719
QY 2281 GGCACGGGCACAGCGGTGAGAGCCCTGCAGGCTTCCCGGCCATCTCTGGAGCCGCTGACC 2340
Db 1720 GGCACGGGCACAGCGGTGAGAGCCCTGCAGGCTTCCCGGCCATCTCTGGAGCCGCTGACC 1779
QY 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 1780 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 1839
QY 2401 CCCAAGAGCCTCGGAGGACAGTCTCTCATCCAGGACCCCGGGGACCCCTCTGTC 2460
Db 1840 CCCAAGAGCCTCGGAGGACAGTCTCTCATCCAGGACCCCGGGGACCCCTCTGTC 1899
QY 2461 TTGCACAACAGCGCTCTCAGCCTCTCCAAACAGGTGGAGCGCGGAGCAGAGGCCCCCT 2520
Db 1900 TTGCACAACAGCGCTCTCAGCCTCTCCAAACAGGTGGAGCGCGGAGCAGAGGCCCCCT 1959
QY 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGCGTGTCTCCGAGAGTGGCGAGTGGCGTGCAG 2580
Db 1960 GCACGCTGGGTGGCTGGCAGCTGGGGCGCGTGTCTCCGAGAGTGGCGAGTGGCGTGCAG 2019
QY 2581 AAGCGGCGGTGGACTGTGGGGCTCCGCGGGCAGCGCAGCTGCCCTGTGTATGCA 2640
```

```
Db 2020 AAGCGGCGGTGGACTGCCGGGCTCCGCCGGCAGCGCTCCCTGCTGTGAIGCA 2079
QY 2641 GCCCATCGGCCCTGGAGACACAAAGCTTGGGGGAGCCCTGCCCACTGGGAGTTCAGC 2700
Db 2080 GCCCATCGGCCCTGGAGACACAAAGCTTGGGGGAGCCCTGCCCACTGGGAGTTCAGC 2139
QY 2701 GCCTGGTCAACCTGTCTCCAAGAGCTGGCGCGGGGATTTCAAGAGCGCTCACTCAAGTGT 2760
Db 2140 GCCTGGTCAACCTGTCTCCAAGAGCTGGCGCGGGGATTTCAAGAGCGCTCACTCAAGTGT 2199
QY 2761 GTGGGCCACGGAGCGCGCTGTGGCCGGGACCAAGTGCACCTTGCAACCAAGCGCCAG 2820
Db 2200 GTGGGCCACGGAGCGCGCTGTGGCCGGGACCAAGTGCACCTTGCAACCAAGCGCCAG 2259
QY 2821 GAGCTGGACTTCTGGCTCTCTGAGGCCGTGTGA 2853
Db 2260 GAGCTGGACTTCTGGCTCTCTGAGGCCGTGTGA 2292
```

Search completed: May 16, 2003, 00:28:37
Job time : 5635 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2003, 22:28:13 ; Search time 1084 Seconds
(without alignments)
12487.358 Million cell updates/sec

Title: US-09-965-631-3

Perfect score: 2853

Sequence: 1 atgtcttctgtggcctcct.....gcgtctgtgagcgtgtcgtga 2853

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 614469 seqs, 2372291704 residues

Total number of hits satisfying chosen parameters: 1228938

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2851.4	99.9	2853	8	US-10-391-364-76
2	2851.4	99.9	2853	8	US-10-391-364-78
3	2851.4	99.9	2853	8	US-10-170-235-10538
4	2848.2	99.8	2853	11	US-60-453-135-7278
5	2848.2	99.8	2853	11	US-60-453-050-7278
6	2817.6	98.8	2930	9	US-10-311-035-32
7	2299.4	80.6	5714	9	US-10-144-771-12354
8	664.8	23.3	3933	10	US-60-455-444-3860
9	664.4	23.3	3711	6	US-09-949-002-108
10	664.4	23.3	4016	6	US-09-724-676A-18640
11	664.4	23.3	4016	6	US-09-724-676A-18640
12	662.8	23.2	2664	6	US-09-949-002-248
13	662.8	23.2	3933	8	US-10-170-235-34404
14	636	22.3	2856	9	US-10-144-771-20257
15	629.8	22.1	4309	1	PCT-US02-24567-57
16	629.8	22.1	4309	9	US-10-210-120-57
17	629.8	22.1	4439	8	US-10-170-235-18710
18	629.8	22.1	4459	9	US-10-159-563-192
19	629.8	22.1	4459	9	US-10-159-563-308
20	629	22.0	4639	10	US-60-455-444-3745
21	629	22.0	4639	11	US-60-452-680-11289
22	629	22.0	4639	11	US-60-453-135-7150

23	629	22.0	4639	11	US-60-453-050-7150	Sequence 7150, Ap
24	628.2	22.0	2853	1	PCT-US02-41225A-18	Sequence 18, Appl
25	628.2	22.0	3430	1	PCT-US02-41225A-17	Sequence 17, Appl
26	628.2	22.0	4471	11	US-60-440-068-133	Sequence 133, App
27	612.8	21.5	4878	9	US-10-152-319A-1840	Sequence 1840, Ap
28	612.8	21.5	4878	9	US-10-301-856-1040	Sequence 1040, Ap
29	612.8	21.5	4878	9	US-10-338-044-2196	Sequence 2196, Ap
30	609.6	21.4	3172	6	US-09-724-676-32233	Sequence 32233, A
31	609.6	21.4	3172	6	US-09-724-676A-32233	Sequence 32233, A
32	585.6	20.5	4305	10	US-60-455-444-1128	Sequence 1128, Ap
33	585.6	20.5	4305	11	US-60-452-680-3792	Sequence 3792, Ap
34	585.6	20.5	4305	11	US-60-453-135-2298	Sequence 2298, Ap
35	585.6	20.5	4305	11	US-60-453-050-2298	Sequence 2298, Ap
36	583.2	20.4	2514	1	PCT-US02-41225A-20	Sequence 20, Appl
37	583.2	20.4	3593	6	US-09-949-002-278	Sequence 278, App
38	583.2	20.4	4192	5	US-09-634-2870-1	Sequence 1, Appl
39	583.2	20.4	4301	1	PCT-US02-41225A-19	Sequence 19, Appl
40	583.2	20.4	4305	8	US-10-170-235-30070	Sequence 30070, A
41	583.2	20.4	4307	6	US-09-949-002-94	Sequence 94, Appl
42	583.2	20.4	4307	8	US-10-358-283-31	Sequence 31, Appl
43	581.6	20.4	4407	8	US-10-015-392A-316	Sequence 316, App
44	581.6	20.4	4407	8	US-10-013-912A-316	Sequence 316, App
45	581.6	20.4	4407	8	US-10-015-394A-316	Sequence 316, App

ALIGNMENTS

RESULT 1

US-10-391-364-76
; Sequence 76, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroli, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MPI03-0190NMIM
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

NAME/KEY: CDS
LOCATION: (1)....(2853)
US-10-391-364-76

Query Match 99.9%; Score 2851.4; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCTTTCTGCTGGGATCCTAACCTTGGCTTTCGCGGGCGAACCGCTGAGGCTCTGAG 60
Db 1 ATCTTTCTGCTGGGATCCTAACCTTGGCTTTCGCGGGCGAACCGCTGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGGCCGCGCTAC 120
Db 61 CCAGAGCGGAGGTAGTCTTCCATCCGACTGGACCCGGACATTAACGGCCGCGCTAC 120
QY 121 TACTGGGGGGTCCGAGGACTCCGGGGATCAGGGACTCATTTTCAGATCACAGCATTT 180
Db 121 TACTGGGGGGTCCGAGGACTCCGGGGATCAGGGACTCATTTTCAGATCACAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGCTTTCAGACCTGGGACGC 300
Db 241 ACTGAGCATCTGGGCTGCCCTCCAGGGGCTCACCGGGGCTTTCAGACCTGGGACGC 300
QY 301 TGCTTTATTTCTGGGACGTGAACCGCGGAGCTCGTTCGCTGTGAGCCCTGTGC 360
Db 301 TGCTTTATTTCTGGGACGTGAACCGCGGAGCTCGTTCGCTGTGAGCCCTGTGC 360
QY 361 GGGGGGCTCCGCGAGCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC 420
Db 361 GGGGGGCTCCGCGAGCTTTGGCTACCGAGGCGCGAGTATGTCATTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGGCGAGCGAAGACGACGAGGCGGACACCTTCTCAGCGCGCG 480
Db 421 AATGCTAGCGCGCGGGGCGAGCGAAGACGACGAGGCGGACACCTTCTCAGCGCGCG 480
QY 481 GGTGTTCCGGCGGGCTTCCGAGAGCCCGACCTCTCGTGGGGGTGCTCGGGCTGG 540
Db 481 GGTGTTCCGGCGGGCTTCCGAGAGCCCGACCTCTCGTGGGGGTGCTCGGGCTGG 540
QY 541 AACCCGCCATCTACGGGCTTGGACCTTACAGCGCGGCGGCGGCGGCTTCGGGGAG 600
Db 541 AACCCGCCATCTACGGGCTTGGACCTTACAGCGCGGCGGCGGCGGCTTCGGGGAG 600
QY 601 AGTCGTAGCGCGGCGAGGCTGGGGCGGCGAAGCTTTGCTGTCTATCCCGGGTACGTG 660
Db 601 AGTCGTAGCGCGGCGAGGCTGGGGCGGCGAAGCTTTGCTGTCTATCCCGGGTACGTG 660
QY 661 GAGAGCTGTGTGCGGGAGAGTCAATGGTCAAGTTCACGCGCGGACCTTGGAAACAT 720
Db 661 GAGAGCTGTGTGCGGGAGAGTCAATGGTCAAGTTCACGCGCGGACCTTGGAAACAT 720
QY 721 TATCTGTGAGCTGTGTGCGAAGCGGCGGCGGCTTACGCGCATCCGAGACTCTCAAC 780
Db 721 TATCTGTGAGCTGTGTGCGAAGCGGCGGCGGCTTACGCGCATCCGAGACTCTCAAC 780
QY 781 CCCATCAACATCTGTGTGCTCAAGGCTCTCTTTAGAGATCGTGAACCTCCGGGCCAAG 840
Db 781 CCCATCAACATCTGTGTGCTCAAGGCTCTCTTTAGAGATCGTGAACCTCCGGGCCAAG 840
QY 841 GTCACCGGCAATGGGGCTTACGCTGCGCAACTTCTGCGCTGGGAGAGAGCTGAAC 900
Db 841 GTCACCGGCAATGGGGCTTACGCTGCGCAACTTCTGCGCTGGGAGAGAGCTGAAC 900
QY 901 AAAGTGTGAGTACAGACCCCGAGTACTGGGACACTGCCATCTCTTCAACAGGCGAGGAC 960
Db 901 AAAGTGTGAGTACAGACCCCGAGTACTGGGACACTGCCATCTCTTCAACAGGCGAGGAC 960
QY 961 CTGTGTGGAGCCACCCTGTGACACCCCTGGGATGGCTGATGTGGGTACCATGTGTGAC 1020

Db 961 CTGTGTGGAGCCACCCTGTGACACCCCTGGCATGTGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGCTGCTCTGTCTATTGAGGAGGATGGGCTTCCATCAGCTTTCACCACTGCC 1080
Db 1021 CCCAAGAGAGCTGCTCTGTCTATTGAGGAGGATGGGCTTCCATCAGCTTTCACCACTGCC 1080
QY 1081 CACGAGCTGGGCCACGTGTTCAACATCCCCCATGACAAATGTAAAGTCTGTGAGGAGTG 1140
Db 1081 CACGAGCTGGGCCACGTGTTCAACATCCCCCATGACAAATGTAAAGTCTGTGAGGAGTG 1140
QY 1141 TTTGGGAGCTCCGAGCAACCCACATGATGTCGCCGACCTCATCCAGATCGACCTGCC 1200
Db 1141 TTTGGGAGCTCCGAGCAACCCACATGATGTCGCCGACCTCATCCAGATCGACCTGCC 1200
QY 1201 AACCCCTGGTCAAGTGTGCTGCTCATCACGACTTCTTGGACAGCGGCGACGCT 1260
Db 1201 AACCCCTGGTCAAGTGTGCTGCTCATCACGACTTCTTGGACAGCGGCGACGCT 1260
QY 1261 GACTGCTCTTGGACCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTGCCGGCGCC 1320
Db 1261 GACTGCTCTTGGACCAACCCAGCAAGCCATCTCCCTGCGCGAGGATCTGCCGGCGCC 1320
QY 1321 AGCTACACCTTGACGACGAGTGGCTTTTGGCTGGGCTCCAAGCCCTGTCTCT 1380
Db 1321 AGCTACACCTTGACGACGAGTGGCTTTTGGCTGGGCTCCAAGCCCTGTCTCTCT 1380
QY 1381 TACTGAGTACTGACCAACCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1440
Db 1381 TACTGAGTACTGACCAACCTGTGTGACCGGGAAGGCCAAGGACAGATGGTGTGC 1440
QY 1441 CAGACCCGCCACTTCCCTTGGGCGGATGGACACAGCTGTGGCGAGGCGCAAGCTCTC 1500
Db 1441 CAGACCCGCCACTTCCCTTGGGCGGATGGACACAGCTGTGGCGAGGCGCAAGCTCTC 1500
QY 1501 AAAGGGGCTCGTGGAGAGACAACTCAACAGCACAGGGTGGATGTTCTCTGGGCC 1560
Db 1501 AAAGGGGCTCGTGGAGAGACAACTCAACAGCACAGGGTGGATGTTCTCTGGGCC 1560
QY 1561 AAATGGGATCCTATGCGCCCTGTCTCGGCGACATGTGTGGGGGCGTGCAGCTGGCAGG 1620
Db 1561 AAATGGGATCCTATGCGCCCTGTCTCGGCGACATGTGTGGGGGCGTGCAGCTGGCAGG 1620
QY 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGCGAGTACTCGGAGGAGTCAAGGCTG 1680
Db 1621 AGGAGTGCACCAACCCACCCCTGCCAACGGGCGAGTACTCGGAGGAGTCAAGGCTG 1680
QY 1681 AAATACCATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAGAGCTTCGG 1740
Db 1681 AAATACCATCTGCAATCTGGAGCCCTGCCAGCTCAGCTCCGGAAGAGCTTCGG 1740
QY 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTTACAACAGCACACCGGCTCAGCTCTGCC 1800
Db 1741 GAGGAGCAGTGTAGGCTTTCAACGGCTTACAACAGCACACCGGCTCAGCTCTGCC 1800
QY 1801 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGGACAGTCAAGCTCATCTGC 1860
Db 1801 GTGGCATGGTGGCCCAAGTACTCCGGCGTGTCTCCCGGGGACAGTCAAGCTCATCTGC 1860
QY 1861 CGAGCAATGGCAGTGGCTTCTATGTGTGGCACCACCAAGGTTGGAGCGGACCGCTG 1920
Db 1861 CGAGCAATGGCAGTGGCTTCTATGTGTGGCACCACCAAGGTTGGAGCGGACCGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCCGCTGTCTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCCGCTGTCTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
QY 1981 GGGAACTGGGCTCCAAAGAGATTCACAAAGTGTGGGTGTGTGGGGGAGACAAAG 2040
Db 1981 GGGAACTGGGCTCCAAAGAGATTCACAAAGTGTGGGTGTGTGGGGGAGACAAAG 2040
QY 2041 AGCTGCAAGAGGCTGACTGGACTCTTCAACAGCCCATGCTACAAATTCGTGTG 2100

Db 2041 AGCTGCAAGAAGGTGACTGGACTCTTACCAAGCCCATGATGGCTACAAATTTCTGGTGTG 2100
QY 2101 GCCATCCCGCAGGCGCTCAGCATCGACATCCGCGAGCGGTTACAAAGGCTGATC 2160
Db 2101 GCCATCCCGCAGGCGCTCAGCATCGACATCCGCGAGCGGTTACAAAGGCTGATC 2160
QY 2161 GGGGATGACAACACTACCTGGCTCTGAAGAACAGCAAGCAAGTACTTGTCTCAACGGGCAT 2220
Db 2161 GGGGATGACAACACTACCTGGCTCTGAAGAACAGCAAGCAAGTACTTGTCTCAACGGGCAT 2220
QY 2221 TTCGTTGGTTCGGGGTGGAGCGGACCTGTGTGAAGGCGAGTCTGCTCGGTACAGC 2280
Db 2221 TTCGTTGGTTCGGGGTGGAGCGGACCTGTGTGAAGGCGAGTCTGCTCGGTACAGC 2280
QY 2281 GGCACGGGCACAGCGGTGAGAGCTGAGGCTTCCCGGCCCATCTGGAGCCGCTGACC 2340
Db 2281 GGCACGGGCACAGCGGTGAGAGCTGAGGCTTCCCGGCCCATCTGGAGCCGCTGACC 2340
QY 2341 GTGGAGGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTCTATCTG 2400
Db 2341 GTGGAGGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGGGAGGACAAAGTCTCTCATCCCAAGAGCCGCCGGGACCCCTCTGTG 2460
Db 2401 CCCAAGAGCCTCGGGAGGACAAAGTCTCTCATCCCAAGAGCCGCCGGGACCCCTCTGTG 2460
QY 2461 TTGCACACAGCGTCTCAGGCTCTCCAAACAGGTGGAGCAGCGGAGCCGACAGCCGCT 2520
Db 2461 TTGCACACAGCGTCTCAGGCTCTCCAAACAGGTGGAGCAGCGGAGCCGACAGCCGCT 2520
QY 2521 GCACGCTGGGTGGCTGGAGCTGGGGCCGTGCTCCGGAGCTGCGGAGTGGGCTGCGAG 2580
Db 2521 GCACGCTGGGTGGCTGGAGCTGGGGCCGTGCTCCGGAGCTGCGGAGTGGGCTGCGAG 2580
QY 2581 AAGCGGGGGTGGAGTGTGCGGGTCCCGCGGGAGCGACGCTGCTGCTGTGTGTGCA 2640
Db 2581 AAGCGGGGGTGGAGTGTGCGGGTCCCGCGGGAGCGACGCTGCTGCTGTGTGTGCA 2640
QY 2641 GCCATCGCGCGTGGAGACACAAAGCCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 2700
Db 2641 GCCATCGCGCGTGGAGACACAAAGCCTGCGGGAGCCCTGCCACCTGGAGCTCAGC 2700
QY 2701 GCCTGGTCAACCTGCTCCAAAGAGCTGCGCGCGGGATTTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGGTCAACCTGCTCCAAAGAGCTGCGCGCGGGATTTCAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGGCGAGGAGCGGCTGTGCGCGGAGCAGTGTGCACTTGCACCGCAAGCCCGAG 2820
Db 2761 GTGGGCGAGGAGCGGCTGTGCGCGGAGCAGTGTGCACTTGCACCGCAAGCCCGAG 2820
QY 2821 GAGCTGGACTTCTGCTGCTGAGGCGGTGCTGA 2853
Db 2821 GAGCTGGACTTCTGCTGCTGAGGCGGTGCTGA 2853

RESULT 2
US-10-391-364-78
; Sequence 78, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Welch, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MP103-0190NMIM
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370

; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2853)
US-10-391-364-78

Query Match 99.9%; Score 2851.4; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCTTAACCCCTGGCTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGATCTTAACCCCTGGCTTCGCGGGGGAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTTCCATCCGACTGACCCGCGGACATTAACGCGCGCGCTAC 120
Db 61 CCAGAGCGGAGGTAGTCTTCCATCCGACTGACCCGCGGACATTAACGCGCGCGCTAC 120
QY 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTTCAGATCACAGCATTT 180
Db 121 TACTGGCGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTTCAGATCACAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACCTGACCCCGGATGCTCAGTTCCTGGCTCCCGCTTCTCC 240
Db 181 CAGGAGGACTTTTACCTACACCTGACCCCGGATGCTCAGTTCCTGGCTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGGCTTTCAGACCTGCGACGC 300
Db 241 ACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCACCGGGGGCTTTCAGACCTGCGACGC 300
QY 301 TGCTTCTATTCTGGGACGTGAACCGCGAGCCGAGCTGTTTCGCTGCTGTGAGCCTGTGC 360
Db 301 TGCTTCTATTCTGGGACGTGAACCGCGAGCCGAGCTGTTTCGCTGCTGTGAGCCTGTGC 360
QY 361 GGGGGGCTCCCGGAGGCTTTGGCTACCGGGGCGGAGTATGTTCATTAGCCGCTGCC 420
Db 361 GGGGGGCTCCCGGAGGCTTTGGCTACCGGGGCGGAGTATGTTCATTAGCCGCTGCC 420
QY 421 AATGCTAGCGCGCGCGGCGAGCGCAACAGCGGCGCACACCTTCTTCAGAGCGCGG 480
Db 421 AATGCTAGCGCGCGCGGCGAGCGCAACAGCGGCGCACACCTTCTTCAGAGCGCGG 480
QY 481 GGTGTTCCGGGCGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
Db 481 GGTGTTCCGGGCGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGCTGG 540
QY 541 AACCCCGCATCTTACCGGGGCTTACAGCCCTTACAGCCCGGGGGGCGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTTACCGGGGCTTACAGCCCTTACAGCCCGGGGGGCGGCTTCGGGGAG 600

Db 541 AACCCGCGCATCTACGGGCCCTGGACCCCTTACAAGCCGCGGGGGGCGCTTCGGGGAG 600
QY 601 AGTCGTAGCCGGCGAGGTGTGGCGCGCCCAAGCGTTTCGTGCTATTCGCCGGGTACGTG 660
Db 601 AGTCGTAGCCGGCGAGGTGTGGCGCGCCCAAGCGTTTCGTGCTATTCGCCGGGTACGTG 660
QY 661 GAGACGCTGGTGGTCCGGGACGAGTCAATGGTCAAGTTTCCACGGCGGGACCTTGGACAT 720
Db 661 GAGACGCTGGTGGTCCGGGACGAGTCAATGGTCAAGTTTCCACGGCGGGACCTTGGACAT 720
QY 721 TATCTGCTGACGCTGCTGGCAAGCGCGCGGCGACTCTACCGCCATCCAGCATCTCTCAAC 780
Db 721 TATCTGCTGACGCTGCTGGCAAGCGCGCGGCGACTCTACCGCCATCCAGCATCTCTCAAC 780
QY 781 CCCATCAACATCGTTGTGGTCAAGGTGCTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 840
Db 781 CCCATCAACATCGTTGTGGTCAAGGTGCTGCTTCTTAGAGATCGTGACTCCGGGCCCAAG 840
QY 841 GTCACCGCAATGGCCCTGACGCTGCGCAACTTCTGTCCCTGGCAGAGAAGAGCTGAAC 900
Db 841 GTCACCGCAATGGCCCTGACGCTGCGCAACTTCTGTCCCTGGCAGAGAAGAGCTGAAC 900
QY 901 AAAGTGAGTGACAGACACCCCGAGTACTGGGACACTGCCATCTCTTACACAGGCAAGGAC 960
Db 901 AAAGTGAGTGACAGACACCCCGAGTACTGGGACACTGCCATCTCTTACACAGGCAAGGAC 960
QY 961 CTGTGTGGAGCCACCCTGTGACACCCCTGGGCACTGGCTGATGGGTACCATGTGTGAC 1020
Db 961 CTGTGTGGAGCCACCCTGTGACACCCCTGGGCACTGGCTGATGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGCTGCTCTGTCATGTAGGAGATGGGCTTCCATCAGCCTTACCACCTGCC 1080
Db 1021 CCCAAGAGAGCTGCTCTGTCATGTAGGAGATGGGCTTCCATCAGCCTTACCACCTGCC 1080
QY 1081 CACGAGCTGGGCCACAGTGTCAACATGCCCATGACAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1081 CACGAGCTGGGCCACAGTGTCAACATGCCCATGACAATGTGAAGTCTGTGAGGAGGTG 1140
QY 1141 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTATCCAGATCGACCGTGGC 1200
Db 1141 TTTGGGAAGCTCCGAGCAACACATGATGTCCCGACCCCTATCCAGATCGACCGTGGC 1200
QY 1201 AACCCCTGGTCAAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 AACCCCTGGTCAAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 GACTGCTCTGACCAACACCCAGACCCCATCTCCCTGCCCCAGGATCTGCCGGGCGCC 1320
Db 1261 GACTGCTCTGACCAACACCCAGACCCCATCTCCCTGCCCCAGGATCTGCCGGGCGCC 1320
QY 1321 AGCTACACCTGAGCAGAGTGGCGAGTGGCTTTTGGGCTGGGCTTCCAGCCCTGTCT 1380
Db 1321 AGCTACACCTGAGCAGAGTGGCGAGTGGCTTTTGGGCTGGGCTTCCAGCCCTGTCT 1380
QY 1381 TACATGAGTACTGACCAAGCTGTGGTGCACCGGAGGCAAGGAGCAATGTGTGC 1440
Db 1381 TACATGAGTACTGACCAAGCTGTGGTGCACCGGAGGCAAGGAGCAATGTGTGTGC 1440
QY 1441 CAGACCCGCGACCTTCCCTGGGCGGATGGCACAGCTGTGGCGAGGCAAGCTTGCCTC 1500
Db 1441 CAGACCCGCGACCTTCCCTGGGCGGATGGCACAGCTGTGGCGAGGCAAGCTTGCCTC 1500
QY 1501 AAAGGGCCCTGCGTGAGAGACACACCTCAACAGCAGAGGATGGTTCCTGGGCG 1560
Db 1501 AAAGGGCCCTGCGTGAGAGACACACCTCAACAGCAGAGGATGGTTCCTGGGCG 1560
QY 1561 AAATGGATCCCTTATGGCCCTGCTCGCGACATGTGTGGGGGCGTGCAGCTGGCCAGG 1620
Db 1561 AAATGGATCCCTTATGGCCCTGCTCGCGACATGTGTGGGGGCGTGCAGCTGGCCAGG 1620
QY 1621 AGGCAGTGACCAACCCCTGCTGCAACGGGGCAAGTACTGCGAGGAGTGGAGGTG 1680
Db 1621 AGGCAGTGACCAACCCCTGCTGCAACGGGGCAAGTACTGCGAGGAGTGGAGGTG 1680

QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCGGAAAGAGCTTCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCGGAAAGAGCTTCGG 1740
QY 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACACAGCAGCAGCGGCTCACTCTGCC 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTCAACGGCTTACAACAGCAGCAGCAGCGGCTCACTCTGCC 1800
QY 1801 GTGGCATGGTGGCCCAAGTACTCCCGGGGACAGTGCAGCTCACTCTGC 1860
Db 1801 GTGGCATGGTGGCCCAAGTACTCCCGGGGACAGTGCAGCTCACTCTGC 1860
QY 1861 CGAGCCAATGGCACTGGCTTCTATGTGTGGCACCACCAAGGTGGTGGACCGCAGCTG 1920
Db 1861 CGAGCCAATGGCACTGGCTTCTATGTGTGGCACCACCAAGGTGGTGGACCGCAGCTG 1920
QY 1921 TGCTCTCTGACTCCACCTCGTCTGTGTCAAGGCAAGTGCATCAAGGCTGGGTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCGTCTGTGTCAAGGCAAGTGCATCAAGGCTGGGTGTGAT 1980
QY 1981 GGGAACTGGGCTCCCAAGAAGAGATTCCACAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGGCTCCCAAGAAGAGATTCCACAAGTGTGGGGTGTGGGGGAGACAATAAG 2040
QY 2041 AGCTCCAAAGAGGTGACTGGACTCTTTCACCAAGCCCATGCTATGGCTACAAATTTCTG 2100
Db 2041 AGCTCCAAAGAGGTGACTGGACTCTTTCACCAAGCCCATGCTATGGCTACAAATTTCTG 2100
QY 2101 GCCATCCCGCAGGGCCCTCAAGCATCGACATCGCCAGCGGGTTACAAGGCTGATC 2160
Db 2101 GCCATCCCGCAGGGCCCTCAAGCATCGACATCGCCAGCGGGTTACAAGGCTGATC 2160
QY 2161 GGGGATGACAACTACTCGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGCAT 2220
Db 2161 GGGGATGACAACTACTCGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAACGGCAT 2220
QY 2221 TTCTGTGTGTGGGGTGGAGCGGACCTGTGTGAAGGCAAGTCTGCTGCGGTACAGC 2280
Db 2221 TTCTGTGTGTGGGGTGGAGCGGACCTGTGTGAAGGCAAGTCTGCTGCGGTACAGC 2280
QY 2281 GGCAGGCGCAGCGGTGGAGAGCTTCCCGGCCCATCTGGAGCGCCTGACC 2340
Db 2281 GGCAGGCGCAGCGGTGGAGAGCTTCCCGGCCCATCTGGAGCGCCTGACC 2340
QY 2341 GTGGAGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 2341 GTGGAGTCTCTCCGTGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCCCTCTGTC 2460
Db 2401 CCCAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCCCTCTGTC 2460
QY 2461 TTGCACACACGCTCTCAGGCTTCCAAACAGGTGGAGAGCGGAGCAGAGGCCCT 2520
Db 2461 TTGCACACACGCTCTCAGGCTTCCAAACAGGTGGAGAGCGGAGCAGAGGCCCT 2520
QY 2521 GCAGCTGGGTGGTGGCAGCTGGGGCGGTGCTCCCGGAGCTGCGGCACTGGGCTGAG 2580
Db 2521 GCAGCTGGGTGGTGGCAGCTGGGGCGGTGCTCCCGGAGCTGCGGCACTGGGCTGAG 2580
QY 2581 AAGCGGCGGTGGACTGTCCGGGGTCCCGGGGAGCGGCACTGGGCTGTGATGCA 2640
Db 2581 AAGCGGCGGTGGACTGTCCGGGGTCCCGGGGAGCGGCACTGGGCTGTGATGCA 2640
QY 2641 GCCCATCGGCGGTGGAGACACAGCTTGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
Db 2641 GCCCATCGGCGGTGGAGACACAGCTTGGGGAGCCCTGCCACCTGGGAGCTCAGC 2700
QY 2701 GCCTGGTCAACCTGCTCCAAAGAGTGGCGGGGATTTTCAGAGGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGGTCAACCTGCTCCAAAGAGTGGCGGGGATTTTCAGAGGCGCTCACTCAAGTGT 2760

QY 2761 GTGGCCACGAGGCGGCTGCTGCGCGGACCAAGTGCACCTTGACACCGCAAGCCCGAG 2820
DB 2761 GTGGCCACGAGGCGGCTGCTGCGCGGACCAAGTGCACCTTGACACCGCAAGCCCGAG 2820
QY 2821 GAGCTGGACTTCTGCTGCTGAGGCGGCTGCTGA 2853
DB 2821 GAGCTGGACTTCTGCTGCTGAGGCGGCTGCTGA 2853

RESULT 3
US-10-170-235-10538
; Sequence 10538, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CLO01380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 10538
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-10538

Query Match 99.98; Score 2851.4; DB 8; Length 2853;
Best Local Similarity 100.08; Pred. No. 0;
Matches 2852; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCCTAACCTGCTTTTCGCGGGCGCAACCGTGTGAGGCTCTGAG 60
DB 1 ATGCTTCTGCTGGGATCCTAACCTGCTTTTCGCGGGCGCAACCGTGTGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTCCCATCCGACTGGACCGGACATTAAGCGCGCGGTAC 120
DB 61 CCAGAGCGGAGGTAGTCTCCCATCCGACTGGACCGGACATTAAGCGCGCGGTAC 120
QY 121 TACTGCGGGGTCCGAGGACTCCGGGATCAGGAGCTCATTTTCAGATACAGCATTT 180
DB 121 TACTGCGGGGTCCGAGGACTCCGGGATCAGGAGCTCATTTTCAGATACAGCATTT 180
QY 181 CAGGAGGACTTTTACCTACACTGACGCGGATGCTCAGTTTCTGGTCCCGCTTCTCC 240
DB 181 CAGGAGGACTTTTACCTACACTGACGCGGATGCTCAGTTTCTGGTCCCGCTTCTCC 240
QY 241 ACTGAGCATCTGGGGCTCCCTCCAGGGGCTACCGGGGGCTCTTACAGCTCGGAGGC 300
DB 241 ACTGAGCATCTGGGGCTCCCTCCAGGGGCTACCGGGGGCTCTTACAGCTCGGAGGC 300
QY 301 TCGTTCTATTCTGGGAGCTGAACCGCGAGCGGACTCGTTCTGCTGTGCTGTGTC 360
DB 301 TCGTTCTATTCTGGGAGCTGAACCGCGAGCGGACTCGTTCTGCTGTGCTGTGTC 360
QY 361 GGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGGATGATGATTTAGCCGCTGCC 420
DB 361 GGGGGCTCCCGGAGCCTTTGGCTACCGAGGCGCGGATGATGATTTAGCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGCGGAGCGGACAGCGGACAGCGGAGCGGACCTTCTCCAGGCGGG 480
DB 421 AATGCTAGCGCGCGGCGGAGCGGACAGCGGAGCGGAGCGGACCTTCTCCAGGCGGG 480
QY 481 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGGTGG 540
DB 481 GGTGTTCCGGGGGCGCTTCCGGAGACCCACCTCTCGCTCGGGGTGGCTCGGGGTGG 540
QY 541 AACCCCGGCATCTACGGGCGCTTACAGCGCGCGGCGGCGGCTTCCGGGAG 600
DB 541 AACCCCGGCATCTACGGGCGCTTACAGCGCGCGGCGGCGGCTTCCGGGAG 600
QY 601 AGTCGTAGCGGCGGAGGTCTGGGCGCGCGCAAGCGCTTTCTGTCTATCCCGGTACGTG 660
DB 601 AGTCGTAGCGGCGGAGGTCTGGGCGCGCGCAAGCGCTTTCTGTCTATCCCGGTACGTG 660

QY 661 GAGACGCTGCTGCTGCGGACGAGTCAATGGTCAAGTTCCACGGCGCGACCTGAAACAT 720
DB 661 GAGACGCTGCTGCTGCGGACGAGTCAATGGTCAAGTTCCACGGCGCGACCTGAAACAT 720
QY 721 TATCTGCTGACGCTGCTGGCAAGCGGCGGCGGACTCTACCGCATCCCGAGCATCTCAAC 780
DB 721 TATCTGCTGACGCTGCTGGCAAGCGGCGGCGGACTCTACCGCATCCCGAGCATCTCAAC 780
QY 781 CCATCAACATCGTTGCTGCTCAAGTGTCTTCTTAGAGATCGTGACTCCGGGCGCAAG 840
DB 781 CCATCAACATCGTTGCTGCTCAAGTGTCTTCTTAGAGATCGTGACTCCGGGCGCAAG 840
QY 841 GTCACCGGCAATCGCGCTGACGCTGCGCAACTTCTGTGCTGCGGAGGAGAGCTGAAC 900
DB 841 GTCACCGGCAATCGCGCTGACGCTGCGCAACTTCTGTGCTGCGGAGGAGAGCTGAAC 900
QY 901 AAAGTGAGTGACAAAGCAACCCCGAGTACTGGGACACTGCCATCTCTTACCAGGAGAGAC 960
DB 901 AAAGTGAGTGACAAAGCAACCCCGAGTACTGGGACACTGCCATCTCTTACCAGGAGAGAC 960
QY 961 CTGCTGGAGCCACCACTGTGACACCCCTGGGCATGGCTGATGGGTACCATGTGTGAC 1020
DB 961 CTGCTGGAGCCACCACTGTGACACCCCTGGGCATGGCTGATGGGTACCATGTGTGAC 1020
QY 1021 CCCAAGAGAGCTGCTCTGCTATTGAGGACGATGGGCTTCCATCAGCCTTACCACCTGCC 1080
DB 1021 CCCAAGAGAGCTGCTCTGCTATTGAGGAGGATGGGCTTCCATCAGCCTTACCACCTGCC 1080
QY 1081 CACGAGTGGGCGACGCTGCTTCAACATGCCCATGCAATGTGAAGTCTGTGAGAGAGTG 1140
DB 1081 CACGAGTGGGCGACGCTGCTTCAACATGCCCATGCAATGTGAAGTCTGTGAGAGAGTG 1140
QY 1141 TTTGGGAGCTCCGAGCCACACATGATGTCCCGACCTCATCAGATCGACCTGCC 1200
DB 1141 TTTGGGAGCTCCGAGCCACACATGATGTCCCGACCTCATCAGATCGACCTGCC 1200
QY 1201 AACCCCTGCTGAGCTGCTGCTCATCATCAGGACTTCTTGGACGCGGACGCT 1260
DB 1201 AACCCCTGCTGAGCTGCTGCTCATCATCAGGACTTCTTGGACGCGGACGCT 1260
QY 1261 GACTGCTCTGAGCAACCCAGAGCCCATCTCCCTGCGCGGAGGATCTGCCGGGCGC 1320
DB 1261 GACTGCTCTGAGCAACCCAGAGCCCATCTCCCTGCGCGGAGGATCTGCCGGGCGC 1320
QY 1321 AGCTACACCTGAGCCAGCAGTGGCTTTGGGCTGGCTCCCAAGCCCTGCTCT 1380
DB 1321 AGCTACACCTGAGCCAGCAGTGGCTTTGGGCTGGCTCCCAAGCCCTGCTCT 1380
QY 1381 TACATGAGTACTGCACCAAGCTGTGGTGACCGGAGGAGGAGGACAGATGCTGTC 1440
DB 1381 TACATGAGTACTGCACCAAGCTGTGGTGACCGGAGGAGGAGGACAGATGCTGTC 1440
QY 1441 CAGACCGCCACTTCCCTGGGCGGATGGACACGCTGTGGCGAGGAGGAGGAGCTGCTC 1500
DB 1441 CAGACCGCCACTTCCCTGGGCGGATGGACACGCTGTGGCGAGGAGGAGGAGCTGCTC 1500
QY 1501 AAAGGGGCTGCGGTGGAGAGACACACCTCAACAGCACAGGAGGATGCTTCTGGGCG 1560
DB 1501 AAAGGGGCTGCGGTGGAGAGACACACCTCAACAGCACAGGAGGATGCTTCTGGGCG 1560
QY 1561 AAATGGGATCCCTATGCGCCCTGCTCGGCACATGTGTGGGGGCGGTGACGCTGGCAGG 1620
DB 1561 AAATGGGATCCCTATGCGCCCTGCTCGGCACATGTGTGGGGGCGGTGACGCTGGCAGG 1620
QY 1621 AGGAGGTGACCAACCCCTGCGCAACGGGGGCAAGTACTGCGAGGAGGAGTGAAGG 1680
DB 1621 AGGAGGTGACCAACCCCTGCGCAACGGGGGCAAGTACTGCGAGGAGGAGTGAAGG 1680
QY 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCCGGAAGAGCTTCCGG 1740
DB 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCAGCTCAGCCTCCGGAAGAGCTTCCGG 1740

```
QY 1741 GAGGACGAGTGTGAGGCTTTCAACGGCTACAAACACAGCACCACCGGCTCACTCTGGCC 1800
Db 1741 GAGGACGAGTGTGAGGCTTTCAACGGCTACAAACACAGCACCACCGGCTCACTCTGGCC 1800
QY 1801 GTGCATGGGTGCCCAAGTACTCCGGGTGTCTCCCGGGACAGTGCAGGCTCACTCTGC 1860
Db 1801 GTGCATGGGTGCCCAAGTACTCCGGGTGTCTCCCGGGACAGTGCAGGCTCACTCTGC 1860
QY 1861 CGAGCCAAATGSCACTGGCTACTTCTATGTCTGTGACCCACCAAGTGTGGACGGCAGCTG 1920
Db 1861 CGAGCCAAATGSCACTGGCTACTTCTATGTCTGTGACCCACCAAGTGTGGACGGCAGCTG 1920
QY 1921 TGCTCTCTGACTCCACTCCGCTGTGTGTCCAAAGGCAAGTGCATCAAGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACTCCGCTGTGTGTCCAAAGGCAAGTGCATCAAGGCTGTGAT 1980
QY 1981 GGGAACTGGGCTCCAAAGAGATTCCGCAAGTGTGGGTCTGTGGGGAGACAATAAG 2040
Db 1981 GGGAACTGGGCTCCAAAGAGATTCCGCAAGTGTGGGTCTGTGGGGAGACAATAAG 2040
QY 2041 AGCTGCAAGAGTGTGACTGGACTCTTCCCAAGCCCATGCTACATTTCTGTGGTG 2100
Db 2041 AGCTGCAAGAGTGTGACTGGACTCTTCCCAAGCCCATGCTACATTTCTGTGGTG 2100
QY 2101 GCCATCCCGGAGCGGCTCAAGATCGACATCCGCGAGCGCGGTTACAAAGGCTGATC 2160
Db 2101 GCCATCCCGGAGCGGCTCAAGATCGACATCCGCGAGCGCGGTTACAAAGGCTGATC 2160
QY 2161 GGGATGACAACACTACCTGGCTCTGAAGACAGCAAGGCAAGTACCTGCTCAACGGGCAT 2220
Db 2161 GGGATGACAACACTACCTGGCTCTGAAGACAGCAAGGCAAGTACCTGCTCAACGGGCAT 2220
QY 2221 TTCGTGTGTGCGGCTGAGAGCGGACCTGCTGTGAAGGGCAGTCTGTGGGTACAGC 2280
Db 2221 TTCGTGTGTGCGGCTGAGAGCGGACCTGCTGTGAAGGGCAGTCTGTGGGTACAGC 2280
QY 2281 GGCACGGCACAGGCTGAGAGCGCTTCAGGCTTCGCGGCCCATCTCGAGCGGCTGACC 2340
Db 2281 GGCACGGCACAGGCTGAGAGCGCTTCAGGCTTCGCGGCCCATCTCGAGCGGCTGACC 2340
QY 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCGCTACTCTTCTATCTG 2400
Db 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACCGCCCGGCTCGCTACTCTTCTATCTG 2400
QY 2401 CCCAAGAGCCTCGGAGGACAGTCTCTATCCCAAGACCCCGGGGACCCCTCTGTC 2460
Db 2401 CCCAAGAGCCTCGGAGGACAGTCTCTATCCCAAGACCCCGGGGACCCCTCTGTC 2460
QY 2461 TTGCACAACAGCTCTCAGCCCTCTCCAAAGGTGGAGCGCCGACACAGGCCCCCT 2520
Db 2461 TTGCACAACAGCTCTCAGCCCTCTCCAAAGGTGGAGCGCCGACACAGGCCCCCT 2520
QY 2521 GCACGTGGGTGGCTGGCAGCTGGGGCCGTGTCTCCGAGGTGCGGCAGTGGCCCTGAC 2580
Db 2521 GCACGTGGGTGGCTGGCAGCTGGGGCCGTGTCTCCGAGGTGCGGCAGTGGCCCTGAC 2580
QY 2581 AAGGGCGGTGACATGTGCGGGCTCCGGGCGCAGCGACAGGTCTCCCTGCTGTGATGA 2640
Db 2581 AAGGGCGGTGACATGTGCGGGCTCCGGGCGCAGCGACAGGTCTCCCTGCTGTGATGA 2640
QY 2641 GCCCATCGGCCCTGGAGACACAAGCTGTGGGGAGCCCTGCCCCACCTGGGAGCTCAGC 2700
Db 2641 GCCCATCGGCCCTGGAGACACAAGCTGTGGGGAGCCCTGCCCCACCTGGGAGCTCAGC 2700
QY 2701 GCCTGTGACCCCTGTCTCAAGAGCTGCGGCCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCCTGTGACCCCTGTCTCAAGAGCTGCGGCCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
QY 2761 GTGGCCACGAGGCGGCTGTGGCCCGGGACCAAGTTCGACCGGACGCGCCAG 2820
Db 2761 GTGGCCACGAGGCGGCTGTGGCCCGGGACCAAGTTCGACCGGACGCGCCAG 2820
QY 2821 GAGCTGGACTTCTGCGCTCTGAGCGCGTGTCTGA 2853
```

```
Db 2821 GAGCTGGACTTCTGCGTCTGAGGCGGTGCTGA 2853
```

RESULT 4

```
US-60-453-135-7278
; Sequence 7278, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7278
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-7278
```

```
Query Match 99.8%; Score 2848.2; DB 11; Length 2853;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2844; Conservative 8; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGCTTCTGCTGGGCATCTAACCTGGCTTTCGCCGGCGAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGCATCTAACCTGGCTTTCGCCGGCGAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGGAGTAGTCTGTTCCATCCGACTGGACCCGGACATTAACGGCGCGCTAC 120
Db 61 CCAGAGCGGGAGTAGTCTGTTCCATCCGACTGGACCCGGACATTAACGGCGCGCTAC 120
QY 121 TACTGGCGGGTCCGGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCAATT 180
Db 121 TACTGGCGGGTCCGGAGGACTCCGGGATCAGGACTCATTTTCAGATCACAGCAATT 180
QY 181 CAGGAGGACTTTTACCTACCTGACGCGGATGCTCAGTTCTTTGGTCCCGGCTCTCC 240
Db 181 CAGGAGGACTTTTACCTACCTGACGCGGATGCTCAGTTCTTTGGTCCCGGCTCTCC 240
QY 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACCGGGGCTCTTCAGACCTCGCAGC 300
Db 241 ACTGAGCATCTGGGCGTCCCTCCAGGGGCTCACCGGGGCTCTTCAGACCTCGCAGC 300
QY 301 TGCTTCTATTCTGGGAGCTGAACCGCGAGCGGACTGCTTCGCTGTGTGAGCCTGTGC 360
Db 301 TGCTTCTATTCTGGGAGCTGAACCGCGAGCGGACTGCTTCGCTGTGTGAGCCTGTGC 360
QY 361 GGGGGCTTCCCGGAGCCTTTGGCTACCGAGCGCGGATGTGCTATAGCCCGCTGCC 420
Db 361 GGGGGCTTCCCGGAGCCTTTGGCTACCGAGCGCGGATGTGCTATAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGGGCGCAGCGCAACAGCAGGGCGCACACCTTCTCCAGCGCGG 480
Db 421 AATGCTAGCGCGCGGGCGCAGCGCAACAGCAGGGCGCACACCTTCTCCAGCGCGG 480
QY 481 GGTGTTCGGGCGGCGCTTCGGGAGACCCACCTCTCGCTCGGGGTGGCTTCGGGCTGG 540
Db 481 GGTGTTCGGGCGGCGCTTCGGGAGACCCACCTCTCGCTCGGGGTGGCTTCGGGCTGG 540
QY 541 AACCCCGCATCTTACGGGCGCTTACAGCGCGGGCGGGCGGGCTTCGGGGAG 600
Db 541 AACCCCGCATCTTACGGGCGCTTACAGCGCGGGCGGGCGGGCTTCGGGGAG 600
QY 601 ACTGCTAGCGCGCGGCTGTGGCGCGCCAGCGTTTCGTGTCTATCCCGCGGTACGTG 660
Db 601 ACTGCTAGCGCGCGGCTGTGGCGCGCCAGCGTTTCGTGTCTATCCCGCGGTACGTG 660
QY 661 GAGACGCTGTGTGCTCGGAGGAGTCAATGCTCAAGTTTCACAGCGCGGACCTTGGAAAT 720
Db 661 GAGACGCTGTGTGCTCGGAGGAGTCAATGCTCAAGTTTCACAGCGCGGACCTTGGAAAT 720
```

Db 661 GAGACGCTGGTGGTCCGGGACGAGTCAATGGTCAAGTTCACCGCCCGGAACTGGAACAT 720
Qy 721 TATCTGTGACGTGTGTGGCAACAGCGCGGAGACTCTACCGCCCATCCCAAGATCTCTCAAC 780
Db 721 TATCTGTGACGTGTGTGGCAACAGCGCGGAGACTCTACCGCCCATCCCAAGATCTCTCAAC 780
Qy 781 CCATCAACATCTGTGTGTCGAAGGTGCTGCTTCTTACAGATCGTCACTCCGGGCCCAAG 840
Db 781 CCATCAACATCTGTGTGTCGAAGGTGCTGCTTCTTACAGATCGTCACTCCGGGCCCAAG 840
Qy 841 GTACCGGCAATCGCGGCTGACGCTGGCAACTCTCTGCTGCGCAGAGAGAGCTGAAC 900
Db 841 GTACCGGCAATCGCGGCTGACGCTGGCAACTCTCTGCTGCGCAGAGAGAGCTGAAC 900
Qy 901 AAAGTGTGACAAAGCAACCCGAGTACTGGACACTGCCATCTCTTACACAGGAGGAC 960
Db 901 AAAGTGTGACAAAGCAACCCGAGTACTGGACACTGCCATCTCTTACACAGGAGGAC 960
Qy 961 CTGTGTGGAGCCACCACTGTGACACCTGGGCAATGGCTGATGGGTACATGTGTGAC 1020
Db 961 CTGTGTGGAGCCACCACTGTGACACCTGGGCAATGGCTGATGGGTACATGTGTGAC 1020
Qy 1021 CCCAAGAGAGCTGCTGTCTGATTTGAGGACCATGGGCTTTCATCAGGCTTCAACACTGCC 1080
Db 1021 CCCAAGAGAGCTGCTGTCTGATTTGAGGACCATGGGCTTTCATCAGGCTTCAACACTGCC 1080
Qy 1081 CACGAGCTGGGCCACCTGTTCACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
Db 1081 CACGAGCTGGGCCACCTGTTCACATGCCCCATGACAAATGTGAAGTCTGTGAGGAGGTG 1140
Qy 1141 TTTGGGAAGCTCCGAGCAACCAATGATGTCCCGGACCTTCATPCAGATCGACCTGGC 1200
Db 1141 TTTGGGAAGCTCCGAGCAACCAATGATGTCCCGGACCTTCATPCAGATCGACCTGGC 1200
Qy 1201 AACCCCTGTGACGCTGAGTGTGCTGATATACAGGCTTCTGTGACAGCGGACAGGT 1260
Db 1201 AACCCCTGTGACGCTGAGTGTGCTGATATACAGGCTTCTGTGACAGCGGACAGGT 1260
Qy 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCATCTCCCTGCTCGAGGATCTGCCGGGCGC 1320
Db 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCATCTCCCTGCTCGAGGATCTGCCGGGCGC 1320
Qy 1321 AGCTACACCTGAGCCAGGAGTGGAGCTGGCTTTGGCGTGGCTCCAGGCTCTGCTCT 1380
Db 1321 AGCTACACCTGAGCCAGGAGTGGAGCTGGCTTTGGCGTGGCTCCAGGCTCTGCTCT 1380
Qy 1381 TACATGCACTACTGCAACAGCTGTGGTGCACCGGGAAGGCAAGGCAAGATGTGTGC 1440
Db 1381 TACATGCACTACTGCAACAGCTGTGGTGCACCGGGAAGGCAAGGCAAGATGTGTGC 1440
Qy 1441 CAGACCCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGCGAGGCTAAGCTCTGCTC 1500
Db 1441 CAGACCCGCCACTTCCCTGGGCGGATGGCACCAGCTGTGCGAGGCTAAGCTCTGCTC 1500
Qy 1501 AAAGGGGCTCGGTGGAGAGACACACCTCAACAGGAGGCTGGATGTTCTCTGGGC 1560
Db 1501 AAAGGGGCTCGGTGGAGAGACACACCTCAACAGGAGGCTGGATGTTCTCTGGGC 1560
Qy 1561 AAATGGGATCCCTATGGCCCTGCTCGGCACATGTGTGCGGCGTGCAGCTGGCCAGG 1620
Db 1561 AAATGGGATCCCTATGGCCCTGCTCGGCACATGTGTGCGGCGTGCAGCTGGCCAGG 1620
Qy 1621 AGGAGTGCACCAACCCAGCTGCGCAACGGGGGAGTACTGCGAGGAGTGGAGGTG 1680
Db 1621 AGGAGTGCACCAACCCAGCTGCGCAACGGGGGAGTACTGCGAGGAGTGGAGGTG 1680
Qy 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCTCCGGGAAGTCTCCGG 1740
Db 1681 AAATACCGATCTGCAATCTGGAGCCCTGCCCGAGCTCAGCTCCGGGAAGTCTCCGG 1740
Qy 1741 GAGGAGCAGTGTGAGGCTTTCAAGGCTTACAACGCTTACAACACACCAACCGGCTCACTCTCGC 1800
Db 1741 GAGGAGCAGTGTGAGGCTTTCAAGGCTTACAACGCTTACAACACACCAACCGGCTCACTCTCGC 1800

Qy 1801 GTGCATGGTGCCCAAGTACTCCGGCGTGTCTCCCGGACAAAGTGCAGCTCATCTGC 1860
Db 1801 GTGCATGGTGCCCAAGTACTCCGGCGTGTCTCCCGGACAAAGTGCAGCTCATCTGC 1860
Qy 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGTGGAGCGCAGCTG 1920
Db 1861 CGAGCCAAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGTGTGGAGCGCAGCTG 1920
Qy 1921 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCCGCTGTGTCCAAGGCAAGTGCATCAAGGCTGGCTGTGAT 1980
Qy 1981 GGGAACTCTGGCTCCCAAGAGAGATTCGCAAGTGTGGGTGTGGGGAGACAAATAG 2040
Db 1981 GGGAACTCTGGCTCCCAAGAGAGATTCGCAAGTGTGGGTGTGGGGAGACAAATAG 2040
Qy 2041 AGCTGCAAGAGGTGACTGGACTCTTCAAGGCCATGCAATGGCTTACAAATTTCTGGTG 2100
Db 2041 AGCTGCAAGAGGTGACTGGACTCTTCAAGGCCATGCAATGGCTTACAAATTTCTGGTG 2100
Qy 2101 GCCATCCCCGAGCGCCTCAAGCATTCGACATCCGCGCAGCGGTTACAAAGGCGCTGATC 2160
Db 2101 GCCATCCCCGAGCGCCTCAAGCATTCGACATTCGCGCAGCGGTTACAAAGGCGCTGAT 2160
Qy 2161 GGGATGACAACTACTGGCTCTGAAGACAGCCAAAGCAAGTACCTGCTCAACGGGCAT 2220
Db 2161 GGGATGACAACTACTGGCTCTGAAGACAGCCAAAGCAAGTACCTGCTCAACGGGCAT 2220
Qy 2221 TTCTGTGTGCTGGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGTCTGTACAGC 2280
Db 2221 TTCTGTGTGCTGGCGGTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGTCTGTACAGC 2280
Qy 2281 GCGACGGGACAGCGGTGGAGAGCTTCGCGGCCCATCTCTGGAGCGCGCTGACC 2340
Db 2281 GCGACGGGACAGCGGTGGAGAGCTTCGCGGCCCATCTCTGGAGCGCGCTGACC 2340
Qy 2341 GTGAGGTCTCTCTCGTGGGGAAGATGACACCGCCCGGCTCCCTACTCTTCTATCTG 2400
Db 2341 GTGAGGTCTCTCTCGTGGGGAAGATGACACCGCCCGGCTCCCTACTCTTCTATCTG 2400
Qy 2401 CCCAAAGAGCTCGGGGAGACAAGTCTCTCATCCAAAGGACCCCGGGGACCTCTCTGTC 2460
Db 2401 CCCAAAGAGCTCGGGGAGACAAGTCTCTCATCCAAAGGACCCCGGGGACCTCTCTGTC 2460
Qy 2461 TTGCACACAGCTCTCAGCTCTCCAAACAGGTGGAGCGGACCGGACAGCGCCCT 2520
Db 2461 TTGCACACAGCTCTCAGCTCTCCAAACAGGTGGAGCGGACCGGACAGCGCCCT 2520
Qy 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGGTCTCCGCGAGCTGGCGAGTGGCGCTGCAG 2580
Db 2521 GCACGCTGGGTGGCTGGCAGCTGGGGCGGTCTCCGCGAGCTGGCGAGTGGCGCTGCAG 2580
Qy 2581 AAGCGGCGGTGGACTGTGGGCTCCGCGGACAGCGGCTCCCTGCTGTGTGATGCA 2640
Db 2581 AAGCGGCGGTGGACTGTGGGCTCCGCGGAGCTCCGCGGAGCGGCTCCCTGCTGTGATGCA 2640
Qy 2641 GCGCATCGGCGGTGGAGACACAAGCTCTCGGGAGCGCTGCCCGACCTGGGAGCTCAGC 2700
Db 2641 GCGCATCGGCGGTGGAGACACAAGCTCTCGGGAGCGCTGCCCGACCTGGGAGCTCAGC 2700
Qy 2701 GCTGTGTCACCTCTCTCCAAGAGCTGGCGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Db 2701 GCTGTGTCACCTCTCTCCAAGAGCTGGCGCGGGGATTTTCAGAGCGCTCACTCAAGTGT 2760
Qy 2761 GTGGGCGAGGAGCGGCTGTGGCCGGGACAGTCACTTGCACCGGACCGCCAG 2820
Db 2761 GTGGGCGAGGAGCGGCTGTGGCCGGGACAGTCACTTGCACCGGACCGCCAG 2820
Qy 2821 GAGCTGGACTTCTCGCTCTGAGGCGGCTGCTGA 2853
Db 2821 GAGCTGGACTTCTCGCTCTGAGGCGGCTGCTGA 2853

```
RESULT 5
US-60-453-050-7278
; Sequence 7278, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7278
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-7278

Query Match          99.8%; Score 2848.2; DB 11; Length 2853;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2844; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ATGCTTCTGCTGGGATCCTAACCTGGCTTTCGCGGGGAGAACGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGATCCTAACCTGGCTTTCGCGGGGAGAACGCTGGAGGCTCTGAG 60

Oy 61 CCAGAGCGGAGGTAGTGTTCATCCGACTGGACCCGAGACATTAAAGCGCCGCTAC 120
Db 61 CCAGAGCGGAGGTAGTGTTCATCCGACTGGACCCGAGACATTAAAGCGCCGCTAC 120

Oy 121 TACTGGCGGGGTCGCGGAGTCCGCGGATCAGGGACTCATTTTTCAGATCACAGCATTT 180
Db 121 TACTGGCGGGTCCGCGGAGTCCGCGGATCAGGGACTCATTTTTCAGATCACAGCATTT 180

Oy 181 CAGGAGGACTTTTACTACACTGACCGGATGCTCAGTTCTTGCTCCGCTCTCC 240
Db 181 CAGGAGGACTTTTACTACACTGACCGGATGCTCAGTTCTTGCTCCGCTCTCC 240

Oy 241 ACTGAGATCTGGGCTGCCCTCCAGGGGCTACCGGGGCTCTAGAGCTCGGACGC 300
Db 241 ACTGAGATCTGGGCTGCCCTCCAGGGGCTACCGGGGCTCTAGAGCTCGGACGC 300

Oy 301 TGCTTCTATTCTGGGAGGTGAAGCGCGGAGCTCGTTCGCTGCTGTGAGCTGTGC 360
Db 301 TGCTTCTATTCTGGGAGGTGAAGCGCGGAGCTCGTTCGCTGCTGTGAGCTGTGC 360

Oy 361 GGGGGCTCCGCGGAGCTTTGGCTACGAGCGCGGAGTATGCTATAGCCGCTGCC 420
Db 361 GGGGGCTCCGCGGAGCTTTGGCTACGAGCGCGGAGTATGCTATAGCCGCTGCC 420

Oy 421 AATGCTAGCGCGCGGCGGAGCGGACGAGCGGAGCGGAGTATGCTATAGCCGCTGCC 480
Db 421 AATGCTAGCGCGCGGCGGAGCGGACGAGCGGAGCGGAGTATGCTATAGCCGCTGCC 480

Oy 481 GGTGTTCGCGGCGGCTTCCGAGAGACCCACCTCTCGTTCGCGGCTGCGCTCGGCTGG 540
Db 481 GGTGTTCGCGGCGGCTTCCGAGAGACCCACCTCTCGTTCGCGGCTGCGCTCGGCTGG 540

Oy 541 AACCCCGCATCTACGGGCTGAGCCCTTACAAGCGCGGCGGCGGCTTCGGGAG 600
Db 541 AACCCCGCATCTACGGGCTGAGCCCTTACAAGCGCGGCGGCGGCTTCGGGAG 600

Oy 601 AGTCGTAGCGCGGAGTCTGGCGGCGGCGGAGCGGCTTCTATCCCGGTAGCTG 660
Db 601 AGTCGTAGCGCGGAGTCTGGCGGCGGCGGAGCGGCTTCTATCCCGGTAGCTG 660

Oy 661 GAGACGCTGGTGTGCGGAGGAGTCAATGTTCAAGTTCACGCGCGGAGCTGGAACAT 720
Db 661 GAGACGCTGGTGTGCGGAGGAGTCAATGTTCAAGTTCACGCGCGGAGCTGGAACAT 720

Oy 721 TATCTGCTGACGCTGCTGGCAAGCGGCGGAGCTCTACCGGCTATCCAGCATCTCAAC 780
Db 721 TATCTGCTGACGCTGCTGGCAAGCGGCGGAGCTCTACCGGCTATCCAGCATCTCAAC 780
```

1801	GTGGCATGGGTGGCCCAAGTACTCCGGCGGTGTCTCCCGGGGACAAGTGCACACTCATCTGC	1861
1861	CGAGCCAAATGGCACTGGCTACTTCTATGTGCTGGCACCCCAAGGTGGTGGAGCGCACGCTG	1920
1861	CGAGCCAAATGGCACTGGCTACTTCTATGTGCTGGCACCCCAAGGTGGTGGAGCGCACGCTG	1920
1921	TGCTCTCCTGACTCCACCTCCGCTGTGTGTCACAGGAAGTGCATCAAGGCTGGCTGTGAT	1980
1921	TGCTCTCCTGACTCCACCTCCGCTGTGTGTCACAGGAAGTGCATCAAGGCTGGCTGTGAT	1980
1981	GGGAACCTGGGCTCCAAGAAAGATTCGACAAAGTGTGGGTGTGTGGGGGAGACAATAAG	2040
1981	GGGAACCTGGGCTCCAAGAAAGATTCGACAAAGTGTGGGTGTGTGGGGGAGACAATAAG	2040
2041	AGCTGCAAGAAGGTGACTGGACTCTTCACCAAGCCCATCATGGCTATTTGGTGGTG	2100
2041	AGCTGCAAGAAGGTGACTGGACTCTTCACCAAGCCCATCATGGCTATTTGGTGGTG	2100
2101	GCCATCCCCCGAGCGGCGCTCAAGCATCGACATCGGCACGGCGGTACAAAGGGCTCATC	2160
2101	GCCATCCCCCGAGCGGCGCTCAAGCATCGACATCGGCACGGCGGTACAAAGGGCTCATC	2160
2161	GGGGATGACAACACTACTGCTCTGAAGAACAGCAAGCAAGTACTGCTCAAGGGCAT	2220
2161	GGGGATGACAACACTACTGCTCTGAAGAACAGCAAGCAAGTACTGCTCAAGGGCAT	2220
2221	TTGCTGTGTCTGGCGTGGAGCGGACCTGGTGTGAAGGGCAGTCTGCTCGGTACAGC	2280
2221	TTGCTGTGTCTGGCGTGGAGCGGACCTGGTGTGAAGGGCAGTCTGCTCGGTACAGC	2280
2281	GGCAGGGCACACGGGTGGAGAGCTCGAGGCTTCCGGGCCCATCTGTGAGCCGCTGACC	2340
2281	GGCAGGGCACACGGGTGGAGAGCTCGAGGCTTCCGGGCCCATCTGTGAGCCGCTGACC	2340
2341	GTGGAGTTCCTCTCCGTGGGGAAGTACACCGCCCGGCTCGGTACTCTTCTATCTG	2400
2341	GTGGAGTTCCTCTCCGTGGGGAAGTACACCGCCCGGCTCGGTACTCTTCTATCTG	2400
2401	CCCAAAGAGCTCGGAGGACAAAGTCTCTCATCCCAAGACCCCGGGGACCTCTGTC	2460
2401	CCCAAAGAGCTCGGAGGACAAAGTCTCTCATCCCAAGACCCCGGGGACCTCTGTC	2460
2461	TTGCACAACAGCTGCTCAGCCTCTCAACCAAGTGGAGAGCGGAGCAAGCCGCT	2520
2461	TTGCACAACAGCTGCTCAGCCTCTCAACCAAGTGGAGAGCGGAGCAAGCCGCT	2520
2521	GCACGCTGGGTGGCTGGCAGCTGGGGGCGCTGCTCCGCGAGTGGGGCASTGGCTCGAG	2580
2521	GCACGCTGGGTGGCTGGCAGCTGGGGGCGCTGCTCCGCGAGTGGGGCASTGGCTCGAG	2580
2581	AAGCGGGCGGTGGACTGCTCGGGGCTCCCGCGGAGCGCAAGTCCCTGCTGTGATGCA	2640
2581	AAGCGGGCGGTGGACTGCTCGGGGCTCCCGCGGAGCGCAAGTCCCTGCTGTGATGCA	2640
2641	GCCCATGGGCCGCTGGAGACACAAGCTTCGGGGAGCGCTTGGCCCACTGGAGGCTCAGC	2700
2641	GCCCATGGGCCGCTGGAGACACAAGCTTCGGGGAGCGCTTGGCCCACTGGAGGCTCAGC	2700
2701	GCTTGCTCACCTGCTCCAAGAGCTGCGGCGGGGATTTTCAGAGGCGGCTCACTCAAGTGT	2760
2701	GCTTGCTCACCTGCTCCAAGAGCTGCGGCGGGGATTTTCAGAGGCGGCTCACTCAAGTGT	2760
2761	GTGGGCCACGAGGCGCGGCTGCTGGCCGGGACCAAGTGCACACTTGCACCGCAAGCCCCAG	2820
2761	GTGGGCCACGAGGCGCGGCTGCTGGCCGGGACCAAGTGCACACTTGCACCGCAAGCCCCAG	2820
2821	GAGCTGGACTTCTGCGTCTGAGGCGCGTGCCTGA	2853
2821	GAGCTGGACTTCTGCGTCTGAGGCGCGTGCCTGA	2853

```

: Sequence 32, Application US/10311035
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: YUE, Henry
: APPLICANT: ELLIOTT, Vicki
: APPLICANT: GANDHI, Ameena R.
: APPLICANT: LAL, Preeti
: APPLICANT: AU-YOUNG, Janice
: APPLICANT: TRIBOULEY, Catherine M.
: APPLICANT: DELEGEANE, Angelo M.
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: NGUYEN, Dannie B.
: APPLICANT: LEE, Ernestine A.
: APPLICANT: HAFALIA, April
: APPLICANT: KHAN, Farrah A.
: APPLICANT: CHAWLA, Narinder K.
: APPLICANT: YAO, Monique G.
: APPLICANT: LU, Dyung Aina M.
: APPLICANT: ARVIZO, Chandra S.
: APPLICANT: TANG, Y. Tom
: APPLICANT: WALSH, Roderick T.
: APPLICANT: AZIMZAI, Yalda
: APPLICANT: LU, Yan
: APPLICANT: RAMKUMAR, Jayalaximi
: APPLICANT: XU, Yuming
: APPLICANT: REDDY, Roopa
: APPLICANT: DAS, Depopriya
: APPLICANT: KEARNEY, Liam
: APPLICANT: KALLICK, Deborah A.
: TITLE OF INVENTION: Proteases
: FILE REFERENCE: PI-0123 PCT
: CURRENT APPLICATION NUMBER: US/10/311,035
: CURRENT FILING DATE: 2002-12-10
: PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,9
: PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
: NUMBER OF SEQ ID NOS: 42
: SOFTWARE: PERL Program
: SEQ ID NO 32
: LENGTH: 2930
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 7473089CBI
: US-10-311-035-32

```

	Query Match	98.8%	Score	2817.6	DB	9	Length	2930
	Best Local Similarity	99.6%	Pred. N.0;					
	Matches	2846;	Conservative	0	Mismatches	4;	Indels	6;
								Gaps
Qy	1	ATGCTTCTGCTGGGATCCTAAACCTTGGCTTTTCGCGGGCGGAACCGCTGTGAGGCGTTTGAG	60					
Db	75	ATGCTTCTGCTGGGATCCTAAACCTTGGCTTTTCGCGGGCGGAACCGCTGTGAGGCGTTTGAG	134					
Qy	61	CGAGACGGGGAGGTAGTGGTTCCCATCCGACTGGACCGCGACATTAACGGCGCGTACT	120					
Db	135	CGAGACGGGGAGGTAGTGGTTCCCATCCGACTGGACCGCGACATTAACGGCGCGTACT	194					
Qy	121	TACTGCGGGGTCCCGAGGACTCCGGGGATCAGGGACTCATTTTTTCAGATCACAGCATTT	180					
Db	195	TACTGCGGGGTCCCGAGGACTCCGGGGATCAGGGACTCATTTTTTCAGATCACAGCATTT	254					
Qy	181	CAGSAGGACTTTTACCTACACTCAGCGCGGATGCTCACTTCTGGCTCCCGCCTTCTCC	240					
Db	255	CAGSAGGACTTTTACCTACACTCAGCGCGGATGCTCACTTCTGGCTCCCGCCTTCTCC	314					
Qy	241	ACTGAGCATCTGGGGTCCCCCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTCCGACGC	300					
Db	315	ACTGAGCATCTGGGGTCCCCCTCCAGGGGCTCACCGGGGGCTCTTCAGACCTCCGACGC	374					
Qy	301	TGCTTCTATTCTGGGGACGTGAACGGCGAGCCGGACTCTGCTGCTGTGTGAGGCGTTGCG	360					
Db	375	TGCTTCTATTCTGGGGACGTGAACGGCGAGCCGGACTCTGCTGCTGTGTGAGGCGTTGCG	434					


```
Db 2595 CCCCCGACGCTGGGTGGCTGGAGCTGGGGCCGTGCTGCGCAGCTGGCAGTGGC 2654
QY 2575 CTGAGAAAGCGGGGGGTGGACTGTGCGGGCTCGCGGGG. AGCGACCGGTGCTGCGCTGT 2634
Db 2655 CTGAGAAAGCGGGGGGTGGACTGGCGGGCTCGCGGGG. AGCGACCGGTGCTGCGCTGT 2714
QY 2635 GATGAGCCCATCGCCCGGTGGAGACACAAGCTCGCGG. AGCCCTGCGCCACCTGAG 2694
Db 2715 GATGAGCCCATCGCCCGGTGGAGACACAAGCTCGCGG. AGCCCTGCGCCACCTGAG 2774
QY 2695 CTCAGCGCTGGTCAACCTGCTCCAAAGAGTGGCGCCGGG. ATTTCAGAGCGCTCACTC 2754
Db 2775 CTCAGCGCTGGTCAACCTGCTCCAAAGAGTGGCGCGG. ATTTCAGAGCGCTCACTC 2834
QY 2755 AAGTGTGTGGCCACGAGGCGCGCTGCTGGCCGGGAC. AGTGAACCTTGGACCGGAAG 2814
Db 2835 AAGTGTGTGGCCACGAGGCGCGCTGCTGGCCGGGAC. AGTGAACCTTGGACCGGAAG 2894
QY 2815 CCCAGAGCTGGACTTCTCGGTCTGAGCGCGTGC 2850
Db 2895 CCCAGAGCTGGACTTCTCGGTCTGAGCGCGTGC 2930

RESULT 7
US-10-144-771-12354
; Sequence 12354, Application US/10144771
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CL001321
; CURRENT APPLICATION NUMBER: US/10/144,771
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ. ID NOS: 47235
; SEQ. ID NO 12354
; LENGTH: 5714
; TYPE: DNA
; ORGANISM: HUMAN
US-10-144-771-12354

Query Match 80.6%; Score 2299.4; DB 9; Length 5714;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 2507; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
Db 1 ATGCTTCTGCTGGGATCTTAACCTGGCTTTCGCGGGGAAACCGCTGGAGGCTCTGAG 60
QY 61 CCAGAGCGGAGGTAGTCTGCCATCCGACTGGACCCGCTACATTAACGCGCGCGGTAC 120
Db 61 CCAGAGCGGAGGTAGTCTGCCATCCGACTGGACCCGCTACATTAACGCGCGCGGTAC 120
QY 121 TACTGGGGGTCCCGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT 180
Db 121 TACAGGAGGTACGAGGACTCCGGGATCAGGAGTCAATTTTCAGATCAGACATTT 180
QY 181 CAGGAGGACTTTTACCTACACCTGACCGCGGATGCTCAGTCTGCTGCTCCGCTTCC 240
Db 181 CAGGAGGACTTTTATCTACACCTGACCGCGGATGCTCAGTCTGCTGCTCCGCTTCC 240
QY 241 ACTGAGCATCTGGCGGTCCCGCTCCAGGGGCTCACCGGGGCTCTTCAGACCTGCGAGCG 300
Db 241 ACTGAGTATCTPAGGTGTCGCGCTCGAGAGGCTCACTGGCAGCTCTACAGCTGGAGCG 300
QY 301 TCTCTTATCTGGGAGCTGACCGCGGAGCTGCTGCTGCTGCTGAGCGTGGC 360
Db 301 TCTCTTATCTGGGAGCTGACCGCGGAGCTGCTGCTGCTGCTGAGCGTGGC 360
QY 361 GGGGGGCTCCCGGAGCTTTTGGCTACCGAGCGCGGAGTATGTCATTTAGCCCGCTGCC 420
Db 361 GGGGGTCTCCCGGAGCTTTTGGCTACCGAGCTGGGAGTATGTCATTTAGCCCGCTGCC 420
QY 421 AATGCTAGCGCGCGCGGCGCGACGCGACAGCCAGGGCGCACACCTTCTCCAGCGCGG 480
```

```
Db 421 RACACACGCGCCAGAGGCGCAGCTCAGCCAGGGCGCACACCTTCTCCAGTGGCG 480
QY 481 GGTGTTCGCGCGGGCCCTTCGGGAGACCCACCTCTCGGTGCGGGGTGCGCTTCGAGTGG 540
Db 481 GGTGTTCGCTGTAGGGCCCTTCGGGAGACCCACCTCTCGTTGCGGGGTGCGCTTCGAGTGG 540
QY 541 AACCCCGCATCTTACGCGCCCTTGACCCCTTACAAGCCGCGGGCGGGCTTCGCGGAG 600
Db 541 AACCCCGCATCTTACGCGCCCTTGAGGGCTCTGAGCCCTTATAAGCCAGCGCGGCGGGAG 600
QY 601 AGTCGTAGCCGCGCAGGTGTGCGCGCGCAAGGTTTGTGTATATCCCGCGGTACGTG 660
Db 601 AGCCACAACCCCGCAGGTGTGCGCGCGCAAGGCTTGTGTATATACACAGGTACGTG 660
QY 661 GAGAGCGTGTGTGCGCGCGAGTCAATGTCAAGTTTCCACGGCGCGACCTTGAACAT 720
Db 661 GAGAGCGTGTGTGCGCGCGAGTCAATGTCAAGTTTCCACGGCGCGATTGGAACAT 720
QY 721 TATCTGTGACGCTGTGTGGCAACGCGCGCGGCTCTACCGCATCCCAAGATCTTCAAC 780
Db 721 TATCTGTGACGCTGTGTGGCAACGCGCGCGGCTCTACCGCATCCCAAGATCTTCAAC 780
QY 781 CCATCAACATCTGTGTGTCAGGCTGTCTTCTTAGAGATCTGACTCCGCGGCGCAAG 840
Db 781 CCATCAACATCTGTGTGTCAGGCTGTCTTCTTAGAGATCTGACTCCGCGGCGCAAG 840
QY 841 GTCACCGCAATGCGGGCCCTGACCTCGGCAACTTCTGTGCGTGGCAGAGAGTGAAC 900
Db 841 GTCACAGCAACGCGGGCCCTGACTCTCGGCAACTTCTGTGCGTGGCAGAGAGTGAAC 900
QY 901 AAAGTGAGTGAAGACACCCGAGTACTGGGACACTGCCATCTCTTACACGAGGAC 960
Db 901 AAAGTGAGCGATAGCACCTTGTACTGGACACAGCCATCTCTTACACGAGGAC 960
QY 961 CTGTGTGAGCGACACCTGTGACACCTGTGGCATGGCTGATGTGGGTACCATGTGTGAC 1020
Db 961 CTATCGGGGGTACCACTGTGACACCTGTGGCATGGCTGATGTGGGTACCATGTGTGAC 1020
QY 1021 CCAAGAGAGCTGCTGTCTATTGAGGAGATGGGGTTCATCAGCTTCTCACTACTGCG 1080
Db 1021 CCAAGAGAGCTGCTGTCTATTGAGGAGATGGGGTTCATCAGCTTCTCACTACTGCG 1080
QY 1081 CAGAGCTGGGCGAGGTGTTCAACATGCCCATGACANTGTGAAGTCTGTGAGGAGTGG 1140
Db 1081 CATGAGCTGGGCGATGTTTCAACATGCCCATGACAACTGAAAGTGTGTGAGGAGTGG 1140
QY 1141 TTTGGGAAGCTCCGAGGCAACACATGATGTCCCGACCTCATCCAGATCGACCTGCG 1200
Db 1141 TTTGGGAAGCTCAGAGCCCAACCATGATGTCTCCGACCTCATCCAGATCGACCTGCG 1200
QY 1201 AACCCCTGGTCAGCTGAGTGTGCCATCATCCAGCTTCTTGGACAGCGCGGACGCT 1260
Db 1201 AACCCCTGGTCAGCTGAGCGCTGCCATTTACTGACTTCTTGGATAGTGGGCTATGT 1260
QY 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCTCTCCCTGCCCGAGGATCTGCTGCGGCG 1320
Db 1261 GACTGCTCTCTGGACCAACCCAGCAAGCCCTCATCACCCTGCTGAGGACCTGCTGCGG 1320
QY 1321 AGCTACACCTCTGAGCGCAGTGGAGTGGCTTTTGGCTGGGCTCCAGGCGCTTCTCT 1380
Db 1321 AGCTACAGTTTGGCGCAGTGGAGTGGCTTTTGGGCTGGGCTTAAAGCGCTTGGCCA 1380
QY 1381 TACATGAGTACTGTCACCAAGCTGTGTGTCACCGGAAAGCCAGGACAGATGTGTGCG 1440
Db 1381 TATATGAGTACTGTACAAAGCTGTGTGTCACCGGAAAGCCAGGAGATGTGTGCG 1440
QY 1441 CAGACCGCGACTTCCCTCTGGCGGATGGCACACAGCTGTGGCGAGGCGAGCTTGTGCTTC 1500
Db 1441 CAGACTCGCGACTTCCCTCTGGCGGATGGCACACAGCTGTGGCGGAAAGCAAGTTCGCTTC 1500
QY 1501 AAAGGGCTGTGTGGAGAGACAACTCAACAGCACAGGCTGTGTGCTTCTTGGGCG 1560
```


Db 1501 AAGGAGCCTGGGTGAGAGACACAAACCAAGTACCGGGTGGACGCTCTTGGGCG 1560
Qy 1561 AATAGGATCCCTATGCGCCCTGCTCGGCACATGTGGTGGGCGGTGACGTGGCCAG 1620
Db 1561 AATGGAGGCCCTACGTTCTGCTCCCGACCTGCGGGTGGGGCGGTGACGTGGCCCG 1620
Qy 1621 AGGAGTGCACCAACCCCTGCTCCCAACGGGGGCAAGTACTGCGAGGAGTGGGGTG 1680
Db 1621 AGGAGTGCACCAACCCCTGCTCCCAACGGGTGAAATACTGCGAGGAGTAAAGTG 1680
Qy 1681 AATACCGATCTGCAATCTGAGCGCTGCCCGAGCTCAACCTCCGGAAGAGCTTCGG 1740
Db 1681 AATACCGATCTGCAATCTGAGCGCTGCCCGAGCTCAACCTCCGGAAGAGCTTCGG 1740
Qy 1741 GAGGAGAGTGGAGCTTTTCAACGGCTTACCAACCAAGCAACCAACCGGCTCACTTCG 1800
Db 1741 GAGGAGAGTGGAGCTTTTCAATGGCTTACCAACCAAGCAACCAACCGGCTCACTTCG 1800
Qy 1801 GTGGCATGGTGCACCAAGTACTCGGGCTGTCTCCCGGCAAGTCAAGCTCATCTGC 1860
Db 1801 GTGGCATGGTGCACCAAGTACTCGGGCTGTCTCCCGGCAAGTCAAGCTCATCTGC 1860
Qy 1861 CGAGCAATGGGACTGGCTTCTATGCTGGCACCACCAAGTGGTGGACGGCTG 1920
Db 1861 CGAGCAATGGGACTGGCTTCTATGCTAGCAGCTAAGTGGTGGATGATGCTG 1920
Qy 1921 TGCTCTCTGACTCCACCTCGCTGTGTCCAAAGCAAGTCAAGGCTGGCTGTGAT 1980
Db 1921 TGCTCTCTGACTCCACCTCGCTGTGTCCAAAGCAAGTCAAGGCTGGCTGTGAT 1980
Qy 1981 GGAACCTGGGCTCCCAAGAGAGATTCGACAAGTGTGGGGTGGTGGGGGAGACAATA 2040
Db 1981 GGAACCTGGGCTCCCAAGAGAGATTTGACAAGTGTGGGGTGGTGGGGGAGACAATA 2040
Qy 2041 AGCTGCAAGAGTGTGAGCTCTTCAAGCAAGCTGATGATGCTACAAATTTCTGGTG 2100
Db 2041 AGCTGCAAGAGTGTGAGCTCTTCAAGCAAGCTGATGATGCTACAAATTTCTGGTG 2100
Qy 2101 GCATCCCGGAGGCGCTCAAGATCGACATCGCAGCGCGGTACAAAGGGCTGATC 2160
Db 2101 GCATCCCGGAGGCGCTCCAGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Qy 2161 GGGATGACAACTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAAGCGGAT 2220
Db 2161 GGGATGACAACTACCTGGCTCTGAAGAACAGCAAGCAAGTACCTGCTCAAGCGGAT 2220
Qy 2221 TTGCTGGTGTGCGGCTGAGCGGACCTGGTGTGAGGCGAGTCTGCTGCGGTACAG 2280
Db 2221 TTGCTGGTGTGCGGCTGAGCGGACCTGGTGTGAGGCGAGTCTGCTGCGGTACAG 2280
Qy 2281 GGCACGGGACAGCGGTGAGAGCGCTGAGGCTTCCCGGCGGCTGAGGCGGCTGAGC 2340
Db 2281 GGCACGGGACAGCGGTGAGAGCGCTGAGGCTTCCCGGCGGCTGAGGCGGCTGAGC 2340
Qy 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACGCGCGCGGCTGCTGCTCTCTATCTG 2400
Db 2341 GTGGAGTCTCTCCGTGGGGAAGATGACACGCGCGGCTGCTGCTCTCTATCTG 2400
Qy 2401 CCCAAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTC 2460
Db 2401 CCCAAAGAGCCTCGGAGGACAAAGTCTCTCATCCCAAGGACCCCGGGGACCTCTGTC 2460
Qy 2461 TTGCACAAAGAGTCTCAGGCTCTCCAAAGAGTGGAGAGCGGAGACAGGCGGCTCT 2520
Db 2461 TTGCACAAAGAGTCTCAGGCTCTCCAAAGAGTGGAGAGCGGAGACAGGCGGCTCT 2520
Qy 2521 GCAGCTGGTGGTGGGAGTGGGGGCGCTGCTCCGCGAGCTGCGGAGTGGGCTGCGAG 2580
Db 2521 GCAGCTGGTGGGAGGAGTGGGGGCGCTGCTCCGCGAGCTGCGGAGTGGGCTGCGAG 2580
Qy 2581 AAGGGGCGGTGAGTGTGGGGCTCCGCGGCGAGCGGCTGCTGCTGATGATGA 2640
Db 2581 AAGGGGCGGTGAGTGTGGGGCTCCGCGGCGAGCGGCTGCTGCTGATGATGA 2640

Qy 2641 GCCATCGGCGCTGGAGACACAAAGCCTGGGGAGCGCTGCCACCTGGGAGCTCAGC 2700
Db 2641 GACCATCGGCCATTGGAGAGAGCGAGCTGTGGGAACCTGCGCAACTTGGAGCTCGGC 2700
Qy 2701 GCTGTGTACCCCTGTCTCAAGAGTGTGGCGCGGGGATTTAGAGGGCTCTCAAGTGT 2760
Db 2701 AACTGTGTCCCTGTCTTAAAGCTGTGGCGGTGATTTAAGAGGGCTCGCTCAAGTGT 2760
Qy 2761 GTGGGCGAGGAGCGGCTGTGTCGCGCGGAGCAGCAAGTGTGACCTGACCGCAAGCCCCAG 2820
Db 2761 GTGGGCGAGGAGCGGCTGTGTCGCGCGGAGCAGCAAGTGTGACCTGACCGCAAGCCCCAG 2820
Qy 2821 GAGCTGGACTTCTGCTGCTGCTGAGCGCGCTGCTGA 2853
Db 2821 GAATTAGACTTCTGCTGCTGAGACCGCTGCTGA 2853

RESULT 8

US-60-455-444-3860

; Sequence 3860, Application US/60455444

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: BEGOVICH, Ann

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001455

; CURRENT APPLICATION NUMBER: US/60/455,444

; NUMBER OF SEQ ID NOS: 50986

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3860

; LENGTH: 3933

; TYPE: DNA

; ORGANISM: Homo sapiens

US-60-455-444-3860

Query Match 23.38; Score 664.8; DB 10; Length 3933;

Best Local Similarity 59.08; Pred. No. 1.7e-133;

Matches 1349; Conservative 7; Mismatches 884; Indels 48; Gaps 11;

Qy 157 CTCAATTTTCAGATCAGACATTTTCAGGAGCATTTTACCTACACCTGACGCGGATGT 216
Db 1208 CTGCGCTCCACCTGTCCGCTTCGGCAAGGCTTCTGCTGCGCTGCGCGCGGAGC 1267
Qy 217 CAGTTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGCGTCCCTCCAGGGCTCACC 276
Db 1268 AGCTTCTTGGCGCGGCTTCAAGATGAGCGGCTCCGGGGCTCCGGCGGCG---ACC 1324
Qy 277 GGGGGCTTTCAGACCTGCGACCTGCTTCTTATTTCTGGGACGTGAACGCGAGCGGAC 336
Db 1325 GGGGGCGAGCGGGGCTGCGGGGCTGCTTCTTCTCCGCGACCGCTGAATGGGAGCGGAG 1384
Qy 337 TCGTTCGCTGCTGAGCGCTGCGGGGCTGCGCGGAGCTTCCGCGAGCGCTTGGCTACCGAGCGC 396
Db 1385 TCCTGCGCGGCTGAGCGCTGCGGGCTGAGCGGCTTCTTCTGCTGAGCGGAG 1444
Qy 397 GAGTATGTATTAGCCCTGCCCAATGCTAGCGCGCGCGGCGGAGCGGAGCGGAG 456
Db 1445 GAGTTTCACTCCAGCGGCGGCGGCGGCTCCCTGCTGAGCGGCGGAGCGGAG 1504
Qy 457 GCGCGACACTTCTCCAGCGCGGGGCTTCCGCGGCGGCTTCCGAGAGCGGAGCGGCTCT 516
Db 1505 CGTGGGGTCCCGCGGAGCGCGCGCTCCCGGAGAGCGGAGTGGGAGTGGAGAGG 1564
Qy 517 CGTGGCGGGTG-GCCTCGGCTGGAAACCGCGCATCTTACGGGCGCTGAGCGCTTACAA 575
Db 1565 GGAGAGGCTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624
Qy 576 GCGCGCGGCGGCGGCTTCCGCGGAGAGTCTGAGCGGCGGAGG---TCTGGCGCGCGC 630
Db 1625 GAAGAGAGGAGAGAGGCGCTAGCGAGCGGCGGCGGCGGCTTGGGGGCGGAGTAGGAGC 1684

631 AACGGTTTCGTCTATCCCGCGGTACGTGGAGACGCTGGTGTCCGGACAGTCAATG 590
1685 AAGCGGTTTGTCTGAGGCGCGCTTCGTGGAGACGCTGGTGGCCGATCGCTCCATG 1744
691 GTCAGTTCACCGCGCGGACCTGGAACATTAATCTGACGCTGTGTGCAACGCGGG 750
1745 GTCGCTTTTACCGGGCGGACCTGGAACATTAATCTGACGCTGTGTGCAACGCG 1804
751 CGACTCTACCGCCATCCAGCATCTTCAACCCCATCAACATCGTTGTGTCAAGTGTG 810
1805 CGAATCTACAGCACCACCGACATCAAGATTTCCATCACTGATGTTGGTAAAGTGTG 1864
811 CTCTTTAGAGATCGTACTCCGGGCGGACGCTACCGGCAATCGCGCCCTGAGCGTGG 870
1865 ATCTAGAGATGAATAATGGGCGGACGAGGTGTCCGCAATGGGGGCTTACACTGCGT 1924
871 AACTTTCTGCTGGCAGAGAGAGCTGAACAAAGTGTGTCACAAAGACACCCGAGTACTG 930
1925 AACTTTCTGCACTGGCAGAGCGGCTTCAACAGCCGACGACCGCCACAGAGACTAY 1984
931 GACACTGCCATCTCTTACACAGCAGGACCTGTGTGG----AGCCACACCTGTGACAC 987
1985 GACAGGCCATCTCTTACACAGCAGGACCTGTGTGGAGAGAGCTGTGTGACAC 2044
988 CTGGCATGGCTGATGTGGGTACCATGTGTGACCCAGAGAGCTGTGTGATGAG 1047
2045 CTGGGTGGCAGACATCGGGACCATTTTGTGACCCCAACAAAGCTGTGTGATCGAG 2104
1048 GACGATGGCTTCCATCAGCTTCCACACTGCCACAGCTGGGCGGACGCTTCAACATG 1107
2105 GATGAGGGCTCCAGGCGGCCACACCTGGCCCATCACTAGGACACGCTCTCAGCATG 2164
1108 CCCATGACATGTGAAGTGTGAGAGGTGTGGAGAGCTCCGAGCCACACCATG 1167
2165 CCCACGACGACTTCAAGCCCTGCACAGGCTCTTGGGGCCATGGGCAAGACACCATG 2224
1168 ATGTCCCGACCTTCATCAGATGACGCTGCCAACCCCTGGTGCAGCTGCAGTGTGSC 1227
2225 ATGGACCGCTGTGCTGCTCCACTGACAGAGCTGCTGCTGCTGCTGCTGCTGCTG 2284
1228 ATCATACCGACTTCTTGGAGCGGGGACGGTGAATGCTGCTGCTGCTGCTGCTGCTG 1287
2285 TATCTCAGAGACTTCTGGAGCGGGGACGAGACTGTCTCTGATGCTGCTGCTGCTG 2344
1288 CCCATCTCCCTGCGGAGGATCTCGCGG----GCCAGCTACACCTGAGCGACAG 1341
2345 GCGCTGCGCCCTCCACAGGCTTCCGGGCTGCGATGGCCCTGTACAGCTGGACAGCAG 2404
1342 TCGAGCTGGCTTTGGCGTGGCTTCCAGCCCTGTCTTACA-----TGCAGTAC 1392
2405 TCGAGCGAGATCTTTGGGCGGATTTCCGCCACTGCCCCAACACCTCTGCTCAGGAGCTC 2464
1393 TGCACCAAGCTGTGTG---CACCGGAGGCCAGGACAGATGGTGTGCCACCGCG 1449
2465 TCGCGCCAGCTTTGTGGCCACTGATGGGCTGAGCCCTGTGCCACCAAGAAATGGC 2524
1450 CACTTCCCTGGCGGATGGACACAGCTGTGGGAGGCAAGCTGCTCTCAAGGGGCG 1509
2525 AGCTGCGCTGGCTGAGCGGACGCGCTGCGGCGCTGGCCACTGTCTGAGAAGGCGAGC 2584
1510 TG-----CGTGGAGAGACAACTCAACAGCAGGCTGGATGGTTCCTGGGCCAAA 1563
2585 TGTCTACCTGAGGAGAGTGGAGAGGCGCCAAAGCCGCTGGAGATGGAGGCTGGGACCG 2644
1564 TGGGATCCCTATGGCCCTGCTCGGCACATGTGTGGGGGCTGCAGCTGCCAGGAGG 1623
2645 TGGGGACCTGGGGAGATGTTCTCGGACCTGTGGAGGAGGATGAGTTCACACCGT 2704
1624 CAGTGACCAACCCACCCCTGCCAACGGGGGCAAGTACTGCGAGGAGCTGAGGCTGAAA 1683
2705 GAGTGCAGAGACCCCGAGGCTCAGATGGAGGAAGATGCTGCTGGTGGAGAGCCAG 2764
1684 TACCGATCTGCAATCTGGAGCCCTGCCCCAGCTCAGCGTCCGGAAAGAGCTTCCGGGAG 1743

2765 TACCAGTCATGCCACGAGGAATGCC-----CCTGACGGAAAGCTTCAGGGAG 2818
1744 GAGCAGTGTGAGGCTTTCAACGGCTACAAACACAGCACCACCGCTCACTCTCAGCGTG 1803
2819 CAGCAGTGTGAGAGTAATAATGCCCTACAAATACACTGAGGAGCGGAATCT---CCTG 2875
1804 GCATGGTGCCCAAGTACTCCGGGCTGTCTCCCGGGACAAAGTGCAGACTCAATCTGCGGA 1863
2876 CAGTGGTGCCCAAGATATGCTGGGGTGTCCCGGGACCGCTGCAAGTGTCTGCGGA 2935
1864 GCAATGGCAGCTGCTACTTCTATGTGTGGACACCAAGGTGGTGGAGCGGACGCTGTGTC 1923
2936 GCCCGGGGAGGAGGAGTTCAAAGTGTTCGAGGCAAGGTGATGATGAGGACCTGTGTG 2995
1924 TCTCTGACTCCACCTCCGCTGTGTCTCAAGCAAGTGCATCAAGGCTGGTGTGATGGG 1983
2996 GGGCGAGAAACTGCGCCATCTGTGTGGTGGCAGTGTGTCAAGCGCGCTGTGACCAT 3055
1984 AACCTGGGCTCCAAAGAGAGATTCGACAAAGTGTGGGTGTGTGGGGAGACAAATAGAGC 2043
3056 GTGGTGGACTCCCTCGGAAGCTGGACAAATGCGGGTGTGTGGGGCAAGGCAACTCC 3115
2044 TGAAGAAGTGAAGTGTCTTCAACAGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2103
3116 TGCAGGAAGTGTCCGGGTCCCTCACCCCAACCAATATGCTGCTGCTGCTGCTGCTGCTG 3175
2104 ATCCCGGAGGCGCTTCAAGCATTCGACATCCGCGAGCGGGTTACAAAGGGCTGATCGG 2163
3176 ATCCAGCTGTGGCTCACTAATATGCTGAGCAGCAGCGGACCCCGGGTGTGAGAAC 3235
2164 GATGACAACTACTGCTCTGAAGAACAGCAGCAAGTACTGCTCAACGGGCAATTC 2223
3236 GATGGAACTACTGCTGCTGAAGCGGCTGATGGCAGTACTGCTCAACGGCAACCTG 3295
2224 GTGGTGTGGCGGTGAGCGGAGCTGTGTGAGGCGAGTGTGCTGCTGCTGCTGCTGCTG 2283
3296 GGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3355
2284 ACGGGACAGCGGTGGAGAGCCTTCCGAGGCTTCCCGCCCATCTGCTGCTGCTGCTGCTG 2343
3356 TCCATGCGCCACTGAGCGCTGCGAGGCTTCCGCGCTTCCGAGAGCTCTGACAGTG 3415
2344 GAGTCTCTCTCCGT---GGGGAAGATGACACCGCCCGGGTCCGCTGCTGCTGCTGCTG 2400
3416 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3475
2401 CCCAAGA 2408
3476 CCTAATGA 3483

RESULT 9
US-09-949-002-108
; Sequence 108, Application US/09949002
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 3711
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(3711)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-108

Query Match 23.3%; Score 664.4; DB 6; Length 3711;

Best Local Similarity 59.2%; Pred. No. 2e-133;

Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;

QY 157 CTCATTTTTCAGATCACAGCATTTTCAGGAGGACTTTTACCTACACCTGACGCGCGATGCT 216
DB 865 CTCGCGCTCCACCTGTCGCGCTTCGCAAGGCTTCGTTGCGCTTCGCGCGCGACGAC 924
QY 217 CAGTTCTTGGCTCCGCGCTTCTCCACTGAGCATCTGGCGCTCCCGCTCCAGGGCTCAAC 276
DB 925 AGCTTCTTGGCGCGGCTTCAGATCGAGCGCTTCGCGCGCTTCGCGCGCGGCG---ACC 981
QY 277 GGGGCTCTTCAGACTCGAGCGCTCTCTCTCTGCGGACGCTGAACGCGAGCGCGAC 336
DB 982 GGGGCGAGCGGGCTTCGCGCGCTTTTTCGCGCACCTGATGCGAGCGCGAG 1041
QY 337 TCCTTCTGCTGCTGAGCTGTCGCGGGGCTTCGCGGAGCTTTGGCTACCGAGGCGCC 396
DB 1042 TCCTTGGCGCGCTGAGCTGTCGCGCGGCTTCGCGGCTTCCTCTCTGAGCGCGGAG 1101
QY 397 GAGTATGCTATGAGCGCTGCCAATGCTAGCGCGCGCGCGAGCGCAAGCGCAG 456
DB 1102 GAGTTCACCATCCAGCGGAGCGCGGGGCTCCCTGGCTAGCGCGCACCGCTCGAG 1161
QY 457 GCGGCACACCTTCTCCAGCGCGGGGTTCGCGGCGGGCTTCGCGGAGACCCACCTCT 516
DB 1162 CGCTGGGGTCCGCGGAGCGCGCGCTTCGCGGAGCGCGAGTGGAGGTGGAGCG 1221
QY 517 CGCTGCGGGGTG-GCTCGGGCTGAACCGCGCATCTACGGCGCTGGACCTTCAAA 575
DB 1222 GGAGCGGTTCAGAGGAGAGAGAGAGACACACAGGAGGAGAGAGAGAGAGAG 1281
QY 576 GCGCGCGCGGGGCTTCGCGGAGAGTCTGAGCGCGCGCAGG-----TCTGGCGCGCC 630
DB 1282 GAAGAGGAGGAGAGAGGCTAGCGAGCGCGCGCGCGCTGGGCGCGCGAGTAGGAC 1341
QY 631 AAGCGTTTCTGCTATCCCGGTTACCTGGAGAGCTGGTGGCGGAGCTGCTGCGGAGTCAATG 690
DB 1342 AAGCGTTTGTGCTGAGCGCGCTTCGTTGAGAGCGCTGCTGGTGGCGGAGTGGCTTCATG 1401
QY 691 GTCAGTTCCAGCGCGGAGCTGGAACATTTATCTGCTGAGCGCTGCTGGCAACGCGCGCG 750
DB 1402 GCTGCTTCTACGGGCGGAGCTGCGAGAACACATCTGAGCTTAATGCTGTGGCGAGCC 1461
QY 751 CGACTCTACCGCATCCAGCATCTCAACCGCATCAACCTGCTGGTGAAGTCTG 810
DB 1462 CGAATCTAAGCACCGGAGCATCAAGAAATTCATCAACCTGATGGTGAAGTCTG 1521
QY 811 CTCTTTAGAGATGCTGACTCGGGCGGAGGTCACCGGCAATCGGCGCTGACGCTGCGC 870
DB 1522 ATCGTAGAGATGAATAATGGGCGGAGGCTGTCGACATATGGGGCTTACACTCGGT 1581
QY 871 AACTTCTGCTGCGAGAGAGCTGACAAAGTGTGAGTGTGACAGACCGCGGACTCTGG 930
DB 1582 AACTTCTGAACCTGGGAGCGGCTTTCAACGAGCGGCGGCGCGGCGGAGAGACATAC 1641
QY 931 GACACTGCTCTCTTCAACGAGGAGCTGTGTG---AGCCACCACTGTGACACC 987
DB 1642 GACACGGCATCTGCTCAACAGACAGAACTCTGTGGCGAGGAGGCTGTGTGACACC 1701
QY 988 CTGGGATGGCTGATGTGGGTACCATGTGTGACCCCAAGAGAACTGCTGTGCTATGAG 1047
DB 1702 CTGGGTGTGCGAGACATCGGGACATTTGTGACCCCAACAAAGAGTGTCTCGGTATCGAG 1761
QY 1048 GAGATGGCTTCCATCAGCTTCCACTGCGGAGGCTGGCGAGCTGTTCACATG: 1107
DB 1762 GATGAGGGCTCCAGCGGCGGCGGAGCTGCGGCGGAGTGGCGGAGCTCTCAGCATG 1821
QY 1108 CCCATGACAAATGTGAAGTCTGTGAGGAGGTGTTTGGGAGGCTCCGAGCGCAACCATG 1167

DB 1822 CCCACGAGCATCCAAAGCCTGCACAGCGCTTTCGGGCGCATGGGCAAGCAGCAGCTG 1881
QY 1168 ATGTCCCGAGCCCTCATCAGATCGACGCTGCCAAACCCCTGGTTCAGCTGCTGAGTGTGCC 1227
DB 1882 ATGGCACCGCTGTGCTCCACTGAACAGAGCTGCCCTGGTCCCTCGACGGCATG 1941
QY 1228 ATCATCACCGACTTCTTGACAGCGGCGACGGTGACTGCTCTCTGGACCAACGAGCAAG 1287
DB 1942 TATCTCACAGAGCTTCTGGAGCGGCGGACGAGCTGTCTCTGGATGCGCTGTGGTGGC 2001
QY 1288 CCCATCTCCCTGCCGAGGATCTCGCGG---GCCAGCTACACCTGAGCGAGCAG 1341
DB 2002 GCGCTGCCCTCCCGACAGGCTTCGCGGCGCATGGCTGTACCACTGGACAGCAGCAG 2061
QY 1342 TCGAGCTGGCTTTTGGCGTGGCTTCCAAAGCCCTGTCTTTACA-----TGCAGTAC 1392
DB 2062 TCGACGAGATCTTTGGCGCGGATTTCCGCGCACTGCCCAACACCTCTGCTCAGGACGTC 2121
QY 1393 TGCACCAAGCTGTGGTG---CACCGGAGGCCAAGGAGAGATGGTGTGCCAGCCGC 1449
DB 2122 TCGCGCCAGCTTTGGTGCACACTGATGGGCTGAGCGCTGTGCCACACAGAAATGGC 2181
QY 1450 CACTTCCCTGGCGGATGCGACAGCTGTGCGAGGCGCAAGCTCTGCTCAAGGGGCC 1509
DB 2182 AGCTTGGCTGGCTGACGCGCACCGCTGCGGGCTTGGCACCTCTGCTCAGAAGCAGC 2241
QY 1510 TG-----CGTGGAGAGACAACTCAACAGACAGGTGGATGTTCTTGGGCGAAA 1563
DB 2242 TGTCTACCTGAGGAGAAAGTGGAGGCGCAAGCGCTGTGATGAGGCTGGGCAACG 2301
QY 1564 TGGGATCTCTATGGCGCTGCTCGCGACATGTGCTGGGCGGTGACGTGGCGCAGGAG 1623
DB 2302 TGGGACCTTGGGAGAAATGTTCTCGGACCTGTGGAGGAGTACATTTTTCACACCT 2361
QY 1624 CAGTGCACCAACCCACCCCTGCCAACGGGCGCAAGTACTGCGAGGAGTGGGTTGAA 1683
DB 2362 GAGTGCAGAGGACCCGAGCCTCAGAAATGGAGAAAGTACTGCTGCGTGGAGAGGCAAG 2421
QY 1684 TACCGATCTGCAATTCGAGCGCTGCCCGAGCTCAGCTCCGGAAGAGCTTCCGGAG 1743
DB 2422 TACAGTATGCCACACGAGAGATGCC---CCTGACGGGAAAGCTTCAGGAG 2475
QY 1744 GAGCAGTGTGAGGCTTTCAACGGCTTCAACACAGCAGCAACCGGCTCAGTCTGCGCGT 1803
DB 2476 CAGCAGTGTGAGAGTATAATGCTTACAAATACACTGACATGGACGGGAAATCT---CCTG 2532
QY 1804 GCATGGTGGCGAGTACTCGCGGCTGCTCCCGGAGCAAGTGCACAGCTCATCTGCCGA 1863
DB 2533 CAGTGGTGGCGAGTATGCTGGGCTGTCGCGGAGCGCTGCAAGTGTGTCGCGCA 2592
QY 1864 GCCAATGGCACTGCTTCTATGCTGCGACCCAGGTGGTGGGAGGAGCAATAGAGC 1923
DB 2593 GCCCGGGGAGGAGGAGTTCAAAGTGTTCGAGCGCAAGGTGATGATGACCCCTGTGT 2652
QY 1924 TCTCTGACTCCACTCCGCTGTGTCAAGGCAAGTGCATCAAGGCTGCTGTGATGG 1983
DB 2653 GGGCCAGAAACACTGGCCATCTGTGCTCGTGGCAGTGTGTCAAGCGCGCTGTGATGAT 2712
QY 1984 AACCTGGGCTCCAGAAAGAGATTCGACAAAGTGTGGGTGTGTGGGAGGAGCAATAGAGC 2043
DB 2713 GTGGTGGACTCGCTCGAAGCTGGACAAATGCGGGGTGTGTGGGCGCAAGCACTCC 2772
QY 2044 TGCAGAGGCTGACTGAGCTTTCACCAAGCCATGATGCTGCTACATTTCTGTGGTGGCC 2103
DB 2773 TGCAGAAAGGCTTCGCGGCTCCCTCACCCCAATATGCTACATGATGATGTCAC 2832
QY 2104 ATCCCGGAGGCGCTCAAGCATGACATCGCGCGCGGTTTACAAAGGCTGATCGGG 2163
DB 2833 ATCCAGCTGTGGCGCACTTAATTTGACGTGAAGCAGCGGAGCCACCGGCTGTGCAGAAC 2892
QY 2164 GATGACAACTACCTGGCTCTGAAGAACAGCAAGGAGTACCTGCTCAACGGGCAATTC 2223

Db 2893 GATGGAACTACCTGGCGCTGAAGACGGCTGATGGGACGCTACCTGCTCAAGGCAACCTG 2952
QY 2224 GTGGTGTGGGGGTTGAGCGGACCTGGTGGTGAAGGGGAGTCTGCTGCTGACAGCGGC 2283
Db 2953 GCCATCTCTGTCATAGACGAGACATCTTGGTGAAGGGGACCATCTGTAATAGAGGGC 3012
QY 2284 ACGGGACACGGGTGGAGAGCTGACGCTTCCGGGCCCATCTGGAGCCGCTGACCGTG 2343
Db 3013 TCCATCGCCACCTTGGAGCGCTGACAGCTTCCGGCCCTTGGCAGAGCTCTGACAGTG 3072
QY 2344 GAGTCTCTTCCGT---GGGGAAGATGACACCGCCCGGGTCCGCTACTCTTCTATCTG 2400
Db 3073 CAGTCTCTGACAGTCCCTGGCGAGGCTTCCCGCCAAAAGTCAATATACACCTTCTTGT 3132
QY 2401 CCCAAAGA 2408
Db 3133 CCTAATGA 3140

RESULT 10

US-09-724-676-18640
; Sequence 18640, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18640
; LENGTH: 4016
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-18640

Query Match 23.3%; Score 664.4; Db 6; Length 4016;

Best Local Similarity 59.2%; Pred. No. 2e-133;

Matches 1354; Conservative 0; Mismatches 886; Indels 48; Gaps 11;

QY 157 CTCATTTTTCAGATCACAGGATTCAGAGGAGCTTTTACCTACACTGACGCGGATGCT 216
Db 865 CTCGGCTCCACCTGTGCGCTTCGGCAAGGGCTTCGTGTGGGCTGAGCGCCGACAC 924
QY 217 CAGTTCTTGGCTCCGCTTCTCCATGAGCATCTGGGCGTCCCGCTCCAGGGGCTACC 276
Db 925 AGCTTCTTGGCGCGGAGTTCAAGATCGAGCGCTCGGGGCTCCGGCGGCG---ACC 981
QY 277 GGGGCTCTTCAGACCTCGGACGCTGCTTCTATTCTGGGAGCTGAACGCGGAGCGAC 336
Db 982 GGGGCGAGCGGGGCTCGCGGCTGTTTTTTTTCGGGACCTGTAATGGGGAGCCGAG 1041
QY 337 TCGTTCTGCTGTGAGCCTGTGGGGGCTCCCGGAGGCTTTGGCTACCGAGGCGCC 396
Db 1042 TCGCTGGCGGGCTCAGCCTGTGCGGGGCTGAGCGGCTCTCTGCTGACGCGGAG 1101
QY 397 GAGTATGTCATTAGCGCGCTGCCAATGCTAGCGCGCGCGCGCGAGCGCAACAGCGAG 456
Db 1102 GAGTTACACCTTCAGCGAGGCGCGGGGCTTCTGCTCAGCGCGCACCGCTGAG 1161
QY 457 GCGCACACCTTCTCCAGCGCGCGGGTGTTCGGGGCGGCGCTTCCGAGACCCACCTCT 516
Db 1162 CGCTGGGTCCCGGAGCGCGCGCTCCCGGAGGACCGAGTGGGAGTGGAGAGC 1221
QY 517 CGCTCGGGGTG-GCCTCGGCTGGAACCGCGCATCTACGGCCCTGACCTTACAA 575
Db 1222 GGAGAGGTTCAGAGGAGGAGGAGACACACGAGACAGCAGGAGGAGAGGCCAA 1281
QY 576 GCGCGGGGCGGGCTTCGGGGAGAGTCTGAGCGCGGAGG------TCTGGGCGGCC 630
Db 1282 GAAGAGGAGGAGAGGCGCTAGCGAGCGCGCGCGCTGGGGGCCACAGTAGGAGC 1341
QY 631 AGCGTTTCTGTTCTATCCCGGCTAGTGGAGACGCTGGTGGTCCGGACGAGTCAATG 690

Db 1342 AAGCGTTTGTCTGAGCGCGCTTCTGGGAGACGCTGCTGGTGGCGATGCTCCATG 1401
QY 691 GTCAGTTTCCAGCGCGGACCTTGGACATATATCTCTGACGCTGCTGCGAACGCGCGC 750
Db 1402 GTCGCTTCTTACGGGGCGGACCTTCGAGAACCATCTCTGACGTTATATGTCTGTGACGCC 1461
QY 751 CGACTCTACCGCCCATCCAGCATCTCAAGCCCATCAACATCTTGTGTGTCAGTTCGCTG 810
Db 1462 CGATTCTACAGACACCCAGCATCAAGATTCATCAACCTGATGTGTGTAAGTGTCTG 1521
QY 811 CTCTTAGAGATCTGACTCGGGGCCCAAGTTCACGGCAATGCCGCCCTGACGCTGGCG 870
Db 1522 ATCTAGTAAATGAAAATGGGGCCAGAGTGTGTCGACAAATGGGGGGGCTTACATCGGT 1581
QY 871 AACTTCTGTGCTGGCAGAGAAAGCTGAACAAAGTGAAGTGAACAGCACCCGAGTACTGG 930
Db 1582 AACTTCTGCAACTGGCAGCGGCTTTCACACGCGAGCGACCGCCACCGAGAGTAC 1641
QY 931 GACTGTCCATCTCTTCCAGCAGCAGGACTGTGG---AGCCACACCTGTACACC 987
Db 1642 GACAGCGCATCTCTGCTCACAGACAGAACTTCTGTGGGAGGAGGGGCTGTGTGACACC 1701
QY 988 CTGGCATGGCTGATGTGGGTACCATGTGACCCCAAGAGAAAGTCTCTCTGTCATTGAG 1047
Db 1702 CTGGGTGTGGCAGACATCGGGACCATTTGTGACCCCAACAAAGTCTCTCGGTGATCGAG 1761
QY 1048 GACGATGGGCTTCCATCAGCTTCCACACTTCCAGCAGCTGGCCAGCTGTGTCAACATG 1107
Db 1762 GATGAGGGGCTCCAGGGCGGCCACCTCGGCCATGAATAGGCGACCTCTCTAGCATG 1821
QY 1108 CCCCATGACAAATGTGAAGTCTGTGAGGAGTGTGTTGGGAAGCTCCGAGCCACACATG 1167
Db 1822 CCCCAGAGGACTCCAAGCCCTGCACACGCTCTTCGGGCCCATGGGCAAGCACCAGTG 1881
QY 1168 ATGTCGCGACCTTCATCCAGATCGACCGTGCACACCCCTGGTGCAGCTGCACTGGTGC 1227
Db 1882 ATGGCAGCGCTGTTCGTCCACTGAACACAGACGCTGCGCTGGTCCCTGCGAGCGCATG 1941
QY 1228 ATCATCACGACTTCTCTGACAGCGGCGGAGTGTGCTCTCTGACCAACCCAGCAAG 1287
Db 1942 TATCTCACAGAGCTTCTGACGGCGGCGGAGACTGTCTCTGATGCCCTGTGTGGC 2001
QY 1288 CCACTCTCCCTCCGAGGATCTGCGGGC-----GCCAGCTACACCTTGACGACAGCAG 1341
Db 2002 GCGCTGCCCTCCCGACAGGCTCCCGGGCGGCGATGGCCCTGTACCAGCTTGGACAGCAG 2061
QY 1342 TCGAGCTGGCTTTTGGGCTGGGTCCCAAGCCCTGTCTTACA-----TGCAAGTAC 1392
Db 2062 TGCAGGAGATCTTTGGCGGATTTCCGCCACTTGCACCAACACCTCTGCTCAGGAGCTC 2121
QY 1393 TGCACCAAGCTGTGTG---CACCGGAGGCGCAAGGACAGATGTTGTGCCAGACCGC 1449
Db 2122 TCGGCCAGCTTTGTGTCACACTGATGGGCTGAGCCCTGTGCCACAGAAAGATGGC 2181
QY 1450 CACTTCCCTGGCGGATGGCACCAGCTGTGGCGAGGCGCAAGCTCTGCTCAAGGGGCGC 1509
Db 2182 AGCTTCCCTGGCTGACGCGACGCGCTGCGGGCTGGGCACTCTGCTCAGAAGGAGCAG 2241
QY 1510 TG-----CGTGGAGACACACCTCAAGACAGAGGTTGATGTTCTCTGGGCCAA 1563
Db 2242 TGTCTACCTGAGAGGAAAGTGGAGGCGCCCAAGCCCTGTGTAGTATGAGGCTGGGCGCC 2301
QY 1564 TGGGATCCCTATGCGCCCTGCTCGCGACATGTGTGGGGGCGCTGCACTGGCGAGAGG 1623
Db 2302 TGGGAGCTTGGGAGATGTTCTCGGACCTGTGAGGAGGAGTACATGTTTTCACACCT 2361
...QY 1624 CAGTGCACCAACCCCGCTGCCAACGGGGGCAAGTACTTCCGAGGAGTGAAGGTCGAA 1683
Db 2362 GAGTGCAAGGAGCGGAGCTCTCAATGAGGAAAGATCTGCTGGGTCCGAGAGGCGCAG 2421
QY 1684 TACCGATCTCTCAATCTGGAGCCCTTGCACGCTCCGCGAAAGAGCTTCCGGGAG 1743


```

Qy  2344  GAGGTCCTCTCCGT---GGGGAAGATGACAGCGCCCGGGGTCTACTCTCTTCVAICTG 2400
Db  3073  CAGCTCCTGACAGTCCCTGGCGGAGGTCTTCCGCCCAAAAGTCAAATACACCTTCVTTGTT 3132
Qy  2401  CCCARAGA 2408
Db  3133  CCTATGA 3140

RESULT 12
US-09-949-002-248
; Sequence 248, Application US/09949002
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 248
; LENGTH: 2664
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-248

```

Query Match	23.2%	Score 662.8	DB 6	Length 2664
Best Local Similarity	59.1%	Prod. No. 4.1e-133		
Matches 1353	Conservative	0	Mismatches 887	Indels 48
QY	157	CTCATTTTTCAGATCACAGCAATTCAGGAGGACTTTTAACTACACCTGACGCGGATGCT	216	
DB	151	CTCGCGCTCCACCTGTCGCGCTTCGGAAGGCTTCGTGTGCTGCGCTGCGCGCCGACGAC	210	
QY	217	CAGTTCTTGCGTCGCGGCTTCTCCACTGAGCAATCTGGCGCTGCCCTTCAGAGGGTTCACC	276	
DB	211	AGGTTCTTAGCGCCGATTTCAAGATTCGAGCGCTTCGGGGCTCCGCGCGGGCG---ACC	267	
QY	277	GGGGGCTCTTCAGACCTCGCACTGCTTCTATTCTGGGAGCTGAACGCCAGCGCGAC	336	
DB	268	GGGGGAGCGGGGCTGCGGGCTGCTTCTTCGCGCACGTGAATGGGAGCGCCGAG	327	
QY	337	TCGTTCTGCTGTGTGAGCTGTGCGGGGGGCTCCGCGAGCGCTTGTGCTACGAGAGCGCC	396	
DB	328	TCGCTGGCGGGGTGAGCTGTGCGCGGGGCTGAGCGGCTCTCTTCGTGTACAGCGGAG	387	
QY	397	GAGTATGTCTATTAGCCGCTGCCAATGTGTAGCGCGCGCGGCGAGCGCAACAGCGAG	456	
DB	388	GAGTTTCAACATCCAGCGCGAGGGCGGGGGCTCCCTGGCTAGCGCGCACCGCTTGAG	447	
QY	457	GGCGCACACCTTCTCCAGCGCGGGGTGTTCGGCGGGGCTTCCGGAGACATCCATCTCT	516	
DB	448	CGCTGGGGTTCGCGCGGAGCCCGCCCTCCGCGGAGAGCCGAGTGGAGGTGAGAGCG	507	
QY	517	CGCTGGGGGTG-GCTCGGCTTGAACCCCGCATCTCTACGGGCGCTTGACCTTACAA	575	
DB	508	GGAGAGGCTCAGAGCAGGAGAGNGAGACACAGGAGGAGACCGGAGGAGGAGACCA	567	
QY	576	GCGCGCGCGCGGGCTTCGGGGGAGAGTCGTGTAGCCGCGCAGG-----TCTTGGGTGAGCT	630	
DB	568	GAAGAGGAGGCAAGAGGCGCTTAGCGAGCGCCACCGCCCTTGGGGGCGACGAGTAAGACC	627	
QY	631	AAGCGTTTCTGCTCTATCCGCGGTACGTGTGAGAGCGTGTGGTCGCGGACGAGTCAATG	690	
DB	628	AAGCGGTTGTGTCTGTGAGGCGCGCTTCGTGTGAGAGACGCTGTGTGTGCCCATGCGTCCATG	687	
QY	691	GTCAAGTTCCAGCGCGGAGCTTGAACATATTCTGTGACGCTGTGTGCAACGCGCGCG	750	

Db 688 GCTGCTTCTACGGGGGCGACCTGCAGAACACATCTGACCTTAATGCTGTGGCAGCC 747
QY 751 CCACTCTACCGCCATCCAGCATCTCAACCCCATCAACATGCTGTGGTCAAGGTCTGTG 810
Db 748 CCAATCTACAGCACCACGATCAAGATTCATCACTGATGTTGTTGTAAGTCTGTG 807
QY 811 CTTCTTAGAGATGCTGACTCCGGGCGCAAGGTCAAGGCAATGCGGCTCTACGCTCGGC 870
Db 808 ATCTAGAGATGAAATATGGGCGCCAGAGGTTCGACAAATGCGGGCTTACACTCGCT 867
QY 871 AACTTCTCTGCTGCGCAGAGAGCTGAACAAGTGAAGCAAGCAACCCGAGTACTGG 930
Db 868 AACTTCTCAACTGGCAGCGGCTTTCAACACGAGCCAGGACC3CCACCCAGACACTAC 927
QY 931 GACATGCTCATCTCTTCAACGAGCAGGACCTGTGTGG---AGCCACCACTGTGACACC 987
Db 928 GACAGCGGCTCTCTGCTCAACAGACAGACTTCTGTGGCAGAGGGCTGTGTGACACC 987
QY 988 CTGGGATGCTGATGTGGGTACATGTGTGACCCAGAGAGTGTGTCTGTCTATTGAG 1047
Db 988 CTGGGTGGCAGACATCGGGGACCTTTGTGACCCCAACAAAAGTGTCTGCTGATGAG 1047
QY 1048 GACGATGGGCTTCCATCAGCTTCAACACTGCCCCAGAGCTGGCAGCTGTTCACATG 1107
Db 1048 GATAGGGCTCCAGGCGGCCACACCTGGGCCATGACTGACGACACTCTCAGCATG 1107
QY 1108 CCCATGACATGTGAAGTCTGTGAGGAGTGTGTGGAGCTCGGAGCCCAACCAATG 1167
Db 1108 CCCAGACGACTTCAAGCCCTGCACACAGAGCTTCTTGGGCCCATGGGCAACCAACG 1167
QY 1168 ATGTCCCGACCTTCATCCAGATGACCGTGTGCAACCCCTGTGTGAGCTGAGTGTGCC 1227
Db 1168 ATGGACCGCTGTCTGCTCCACTGAACACGAGCTGCGCTGTGCTGCTGAGTGTGCG 1227
QY 1228 ATATCACCAGCTTCTTGGCAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1287
Db 1228 TATCTCAGAGCTTCTGAGCGGCGGACGAGACTGTCTCTGATGATGCTGCTGCTGCT 1287
QY 1288 CCACTCTCTGCGGAGATGTGCGGGG---GCCAGTACACCTGAGCAGCAG 1341
Db 1288 GCGCTGCGCTTCCCGACAGGCTTCCGGGCGCATGCGGCTGTACCACTGAGCAGCAG 1347
QY 1342 TCGAGCTGCTGCTTGGCGTGGCTCCAAGCCCTGTCTTACA-----TGCACTAC 1392
Db 1348 TGCAGCAGATCTTTGGCGGATTTCCGCACTGCGCCCAACACCTCTGCTCAGAGCTC 1407
QY 1393 TGCACCAAGCTCTGTG---CACGGAGGCAAGGACAGATGCTGTGCTCAGACCCG 1449
Db 1408 TGGCGCAGCTTTGTGCTCAGCTGTGAGGCTGAGCCCTGTGCCACACAGAAATGGC 1467
QY 1450 CACTTCCCTGGCGGATGACACAGCTGTGGGAGGCAAGCTCTGCTCAAGGGGCC 1509
Db 1468 AGCTGCGCTGGCTGAGCGACGCGCTGCGGCTTGGGCACTCTGCTCAGAGGCGAGC 1527
QY 1510 TG-----CGTGGAGAGACAACTCAACAAGCAGGCTGATGTTCTTGGGCCAAA 1563
Db 1528 TGTCTACCTGAGGAGAGTGGAGAGGCCCAAGCCGCTGGCAGATGGAGGCTGGCAGCG 1587
QY 1564 TGGATCTCTATGCGCTCTGCGCAGATGTGTGGGGGCTGACGTGCGCAGGAGG 1623
Db 1588 TGGGACCTTGGGAGAGATGTTCTCGACCTGTGGAGGAGGATACAGTTTTCACACCGT 1647
QY 1624 CAGTGCACCAACCCACCTGCGCAACGCGGCAAGTACTGCGAGGAGTGAAGGTGAAA 1683
Db 1648 GAGTGAAGGACCCGAGCCTCAGATGGAGGAAGTACTGCTGGTGGGAGGCAAG 1707
QY 1684 TACCATCTGCAATCTGGGCTGCGCCAGCTCAGCTCGGCAAGAGCTTGGGAG 1743
Db 1708 TACCATCTGCAATCTGGGCTGCGCCAGCTCAGCTCGGCAAGAGCTTGGGAG 1761
QY 1744 GAGCAGTGTGAGGCTTTCAAGGCTTACAAACAGCAGCAACCGCTCACTCTCGCGCTG 1803
Db 1762 CAGCAGTGTGAGAGTATAATGCTTACATTTACATGACATGGAGGGAATCT---CCTG 1818

QY 1804 GCATGGGTGCCAAGTACTCCGGCGTGTCTCCCGGAGCAAGTCAAGCTCACTTCCGCA 1863
Db 1819 CAGTGGTCCCAAGTATGCTGGGTGTCCCCCGGAGCCGCTGCAAGTTGTCTTCCGCA 1878
QY 1864 GCAATGGCAGTGGCTACTTCTATGTCTGCACCCCAAGTGTGGACGCCAGCTGTGTC 1923
Db 1879 GCGCGGGGAGGAGGAGTCAAGTGTTCGAGCCCAAGGTGATTGATGCAACCTTGT 1938
QY 1924 TCTCTGACTTCCACACTCCGCTGTGTCCCAAGCAAGTGCATCAAGGCTGTGTATGGG 1983
Db 1939 GGGCAGAAACACTGGCCATCTGTCTGCTGCGCAGTGTGTCAAGGCGGCTGTGACCAT 1998
QY 1984 AACTGGGCTTCAAGAAGAGATTCCACAAGTGTGGGTGTGTGGGGAGACAATAGAGC 2043
Db 1999 GTGGTGGACTGCTGCGGAGCTGCAAAATGCGGGGTGTGTGGGGCAAGCAACTCC 2058
QY 2044 TGCAGAAGGTGACTGGACTCTTCCAAAGCCCAATGATGCTACAAATTTCTGTGTGGCC 2103
Db 2059 TGCAGGAAGGTCTCGGGTCCCTCACCCCACTAATTTATGGCTACAAATGACA?GTCAAC 2118
QY 2104 ATCCCGCAGCGGCTCAAGCATCGACATCGCCGAGCGGCTTACAAAGGCTGTATCGGG 2163
Db 2119 ATCCAGCTGTGCTCCACTAATTTGACCTGAAGCAGCGGAGCCCGGCTGTGAGAAC 2178
QY 2164 GATCACAACCTACTGCTCTGAAGAACAGCAAGCAAGTACCTGTCAACGGGCAATTC 2223
Db 2179 GATGGGAACTACTGCGCTGAGACGGCTGATGGCAGTACCTGTCAACGGCAACTG 2238
QY 2224 GTGTGTGCGGCTGGAGCGGACCTGTGTGTGAAGGCGAGTCTGTGCGGTACAGCGGC 2283
Db 2239 GCACTCTCTGCCATAGACGAGGACATCTTGTGAAGGAGGACCATCTGAAAGTACAGCGC 2298
QY 2284 ACGGGACAGCGGTGAGAGCCTGAGAGCTTCCGGGCGCATCTGGAGCGCTGACCGTG 2343
Db 2299 TCCATCGCACCTCGAGCGCTGACAGCTTCCGGCTTGGCAGAGCTCTGACAGTG 2358
QY 2344 GAGTCTCTCTCGT---GGGGAAGATGACAGCGCGCGGTGCGCTACTCTCTATCTG 2400
Db 2359 CAGTCTCTGACAGTCTCTGCGAGGCTTCCCGCCCAAAAGTCAAAATACACTCTTGT 2418
QY 2401 CCCAAGA 2408
Db 2419 CCTAATGA 2426

RESULT 13

US-10-170-235-34404

; Sequence 34404, Application US/10170235

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU

; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

; FILE REFERENCE: CL001380

; CURRENT APPLICATION NUMBER: US/10/170.235

; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514

; SEQ ID NO 34404

; LENGTH: 3933

; TYPE: DNA

; ORGANISM: HUMAN

US-10-170-235-34404

Query Match 23.2%; Score 662.8; DB 8; Length 3933;

Best Local Similarity 59.1%; Pred. No. 4.5e-133;

Matches 1353; Conservative 0; Mismatches 887; Indels 48; Gaps 11;

QY 157 CTCAATTTTCAGATCAGACATTTTCAGGAGGACTTTTACCTACACCTGACGCGGATGCT 216

Db 1208 CTCGGCTCCACCTGTCCCGCTCGGAGGCTTCGTGCTGCGCTTGGCGCCCGGAGAC 1267

QY 217 CAGTCTTGGCTCCCGCTTCTCCACTGAGCATCTGGGCGTCCCGCTCCAGGGGCTCAC 276

Db 1268 AGCTTCCTAGCGCCGAGTTTCAAGATCGAGCGCTCGGGGGCTCGGCCGCGG---ACC 1324
Qy 277 GGGGGCTTTCAGACCTGCGACCTGCTTCTATTCTGGGGAGGTGAACGCGAGCCCGGAC 336
Db 1325 GGGGGGAGGGGGCTGCGCGGCTGCTTCTCTCGGACCGGTGAATGGGAGCCCGAG 1384
Qy 337 TGGTTCGCTGCTGAGCCCTGTGCGGGGGCTCGGGAGGCTTGGCTACGAGGGGCC 396
Db 1385 TCGCTGGCGGGGTGAGCCCTGTGCGGGGCTGAGCGGCTCTCTCTGCTGAGCGGAG 1444
Qy 397 GAGTATGCTTATAGCCGCTGCCAATGCTAGCGCGCGGGGCGAGCGCAACAGCCAG 456
Db 1445 GAGTTCACCATCAGCGCGAGGGCGGGGGGCTCCCTGGCTCAGCGCGCCGCTGCAG 1504
Qy 457 GCGGCACACCTTCTCAGCGCGCGGGGTGTTCGGGGGGGCTTCGCGAGACCCACCTCT 516
Db 1505 CGCTGGGGTCCCGCGAGGCCCGCCCTCCCGGAGGACCGAGTGGGAGTGGAGCG 1564
Qy 517 CGCTCGGGGTG-GCCTCGGGCTGGAAACCGCGCATCTACGGGCGCTGGACCTTACAA 575
Db 1565 GGAGAGGTGAGAGGAGGAGAGGAGACCACAGGAGGACGCGAGGAGGAGCCAA 1624
Qy 576 GCCGGCGGGCGGCTTCGGGAGAGTGTAGCGCGGCGAGG-----TCTGGGCGGCC 630
Db 1625 GAAGAGGAGGAGAGGCGCTAGCGAGCGCGCACCGCCCTGGGGCGCACGAGTAGGACC 1684
Qy 631 AAGCGTTTCGTGCTATCCCGCGGTACGTGGAGCGCTGGTGGTCCGCGACGAGTCAATG 690
Db 1685 AAGCGGTTGTGCTGAGGCGGCTTCTGGGAGCGCTGCTGGTGGCGGATGCGTCCATG 1744
Qy 691 GTCAGTTTCCAGCGCGGACCTGGAAACATATTCTGCTGACGCTCTGCAACGCGGCG 750
Db 1745 GCTGCTTCTAGCGGCGGACCTGCGAGACCATCTGACCTTAATGCTGTGGAGCC 1804
Qy 751 GCACTCTACCGCCATCCAGATCTCTAACCCCATCAACATCTTGTGGTCAAGTGTG 810
Db 1805 CGAATCTACAGACACCCGACGATCAAGAAATTCATCAACCTGATGGTGTAAAAGTGTG 1864
Qy 811 CTTCTTAGAGATCTGACTCGGGGCGCAAGGTCACCGGCAATCGCGGCTTCAGCGTGGC 870
Db 1865 ATCGTAGAAGTAAATGGGCCAGAGGTGTCCGACAATGGGGCTTACACTCGT 1924
Qy 871 AACTTCTGCTGCTGGAGAGAGCTGAACAAAGTGAAGTGAACAGCACCGGAGTACTGG 930
Db 1925 AACTTCTGCAACTGGGAGCGGCTTCAACAGCGCGGAGCGCCACCGCAGAGCATAC 1984
Qy 931 GACACTGCCATCTTTCACAGCGAGGACCTGTGTGG---AGCCACCACTGTGACACC 987
Db 1985 GACAGGCGCATCTGCTCACCAGACAGAACTTCTGTGGCAGGAGGGCTGTGTGACACC 2044
Qy 988 CTGGCATGCTGATGTGGGTACCATGTGTGACCCCAAGAGAGCTGCTGTCTATTGAG 1047
Db 2045 CTGGGTGGCAGACATCGGAGCATTTGTGACCCCAACAAAGCTGTCTGTGATCGAG 2104
Qy 1048 GAGATGGGCTTCATCAGCCTTACACATGCCCCAGAGCTGGGCGACGTGTTCACATG 1107
Db 2105 GATGAGGGCTTCCAGGGGGCCACACCTGGCCCATGAATAGGCGACGTCTCAGCATG 2164
Qy 1108 CCCCATGACATGTGAAGTCTGTGAGGAGGTGTTTGGGAAGCTCCGAGCCCAACACATG 1167
Db 2165 CCCCAGAGGACTCCAGGCTGCAACAGGCTCTCTCGGCCCATGGGCAACCCACATG 2224
Qy 1168 ATGTCCCGACCCCTCATCCAGATGAGCGGTGCCAACCCCTGTGAGCTGCTGCTGTC 1227
Db 2225 ATGGACCGGCTTCTGCTCCACTGAACAGACGCTGCTGCTGCTGCTGCGGCGCATG 2284
Qy 1228 ATCATCAGGACTTCTTGGAGCGGCGGAGTGTGCTGCTGCTGAGCAACCCAGCAAG 1287
Db 2285 TATCTCAGAGAGTCTTGGAGGGGCGGACGAGACTGCTCTGATGCTCCCTGCTGCG 2344
Qy 1288 CCATCTCCTGCGGAGGATCTGCGGGG-----GCCAGTACACCTGAGCCAGAG 1341
Db 2345 GCGCTGCGCCCTCCCGACAGGCGTCCCGGGCGCATGGCGCTGTACCAAGTGGAGCAG 2404

Qy 1342 TCGAGCTGCTTTTGGCGTGGGCTCCAAAGCCCTGTCTTAC-----TGCAGTAC 1392
Db 2405 TCGAGGAGATCTTTGGCGCGGATTTCCGCCACTCCGCCAACACCTCTGCTCAGGAGTC 2464
Qy 1393 TGCACCAAGCTGTGGTGTG---CACCAGGAAAGCCAAAGGACAGATGGTGTGCCAGACCCG 1449
Db 2465 TGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2524
Qy 1450 CACTTCCCTTGGCGGATGCGACCACTGCTGGCAGGAGCAAGCTGCTGCTCAAGAGGGCC 1509
Db 2525 AGCTTCCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2584
Qy 1510 TG-----CGTGGAGAGACACCTCAACAGCAGACAGGTTGGATGTTCTTGGGCGCAA 1563
Db 2585 TGTCTAGCTGAGGAGGAGTGGAGAGCCCAAGCCCGTGGCAGATGGAGGCTGGGACCG 2644
Qy 1564 TGGATCCCTATGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1623
Db 2645 TGGGACCTTGGGAGAAATGTTCTCGGACCTGTGGAGGAGGAGTACAGTTTTCACACCGT 2704
Qy 1624 CAGTGCACCAACCCACCCCTGCTCAACGGGCAAGTACTCGAGGAGTGAAGGTGAA 1683
Db 2705 GAGTGCAGAGACCCGAGCCCTCAGATGGAGGAGATGCTGCTGGTGGAGGAGCAAG 2764
Qy 1684 TACCGATCCTGCAATCTGAGAGCCCTGCCACGCTCAGCCTCCGGAAGAGCTTCCGGAG 1743
Db 2765 TACCAGTATGCCACACGAGGAAATGCCG-----CCTGACGGGAAAGCTTCCAGGAG 2818
Qy 1744 GAGCAGTGTGAGGCTTCAACGGCTACAAACAGCACCAACCGGCTCACCTCTGCGCGTG 1803
Db 2819 CAGCAGTGTGAGAGTATATGCTACATATACATGACATGAGCGGGAATCT---CCTG 2875
Qy 1804 GCATGGGTGCCAAGTACTCCGGGCTGTCTCCCGGAGACAGTGAAGCTCATCTGCGGA 1863
Db 2876 CAGTGGTCCCAAGTATGCTGGGCTGTCCCGCGGAGCGCTGCAAGTCTTCTGCGGA 2935
Qy 1864 GCAATGGCACTGCTACTTCTATGTGCTGGCACCAAGTGGTGGAGCGGAGCTGCTGTC 1923
Db 2936 GCCCGGGGAGGAGCGAGTTCAAGTGTTCGAGGCAAGTGTGATGTCACCTGTGT 2995
Qy 1924 TCTCTGACTCCACTCCCTGCTGTCCAAGCAAGTGCATCAAGCTGCTGTGATGGG 1983
Db 2996 GGGCAGAAAGACTGGGCTTCTGTGCTGGCAGTGTGCTGAGGCGGCTGTGACCAT 3055
Qy 1984 AACCTGGCTCCAGAGAGATTCGACAAAGTGTGGGGTGTGGGGGAGACAAAGAGC 2043
Db 3056 GTGGTGGACTCGGCTCGGAAAGCTGGACAAATGGGGGTGTGGGGGCAAGGCAACTCC 3115
Qy 2044 TGAAGAGGTGACTGGACTCTTCAACAGCCCATGTCATGGCTACAATTTGCTGGTGGCC 2103
Db 3116 TCGAGGAAGTCTCCGGGCTCCCTCACCCCAACCAATATGGCTACAATGACATGTCAAC 3175
Qy 2104 ATCCCGGAGGCGCTCAAGCATCGACATCCGCGAGCGGTTTACAAGGCTGATGCGG 2163
Db 3176 ATCCAGCTGTGCGCACTAATATTGAGTGAAGCAGGAGGAGCCCGGGGTGTGCAAGAC 3235
Qy 2164 GATGACAACTACTGCTGTGAAGAACAGCAAGGCAAGTACTGCTCAAGCGGCAATTC 2223
Db 3236 GATGGAACTACTGCTGGGCTGAAGACGGCTGATGGGCACTGCTGCTCAAGCGCAACTG 3295
Qy 2224 GTGGTGTGGGCTGAGCGGAGCTGTGGTGAAGGCGAGTCTGCTGCGGTACAGGCGC 2283
Db 3296 GCAATCTCTGCAATAGAGCAGGACATCTTGTGAAGAGGAGCACTCTGAAGTACAGGCGC 3355
Qy 2284 ACGGCAGCAGCGTGGAGAGCTTCAGAGGCTTCCCGGCCCATCTCTGGAGCGCTGACGCTG 2343
Db 3356 TCCATCGCACCTTGGAGCGCTTCGAGAGCTTCGCGCCCTTCCAGAGGCTCTGACAGT 3415
Qy 2344 GAGTCTCTCCGT---GGGGAAGATGACACCGCCCGGCTCGCTACTCTCTATCTG 2400
Db 3416 CAGCTCTCTGAGAGTCCCTGGGAGGCTTTCCTCCCAAGTCAATACACCTCTTCTTGT 3475

Db 2500 AACGCCATTC 2509

RESULT 15

PCT-US02-24567-57

; Sequence 57, Application PC/TUS0224567

; GENERAL INFORMATION:

; APPLICANT: Chinnaiyan, Arul M.

; APPLICANT: Rubini, Mark A.

; APPLICANT: Sreekumar, Arun

; TITLE OF INVENTION: Expression Profile of Prostate Cancer

; FILE REFERENCE: UM-07221

; CURRENT APPLICATION NUMBER: PCT/US02/24567

; CURRENT FILING DATE: 2002-12-23

; PRIOR APPLICATION NUMBER: US 60/309,581

; PRIOR FILING DATE: 2001-08-02

; PRIOR APPLICATION NUMBER: US 60/334,468

; PRIOR FILING DATE: 2001-11-15

; NUMBER OF SEQ ID NOS: 123

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 57

; LENGTH: 4309

; TYPE: DNA

; ORGANISM: Homo sapiens

PCT-US02-24567-57

Query Match 22.1%; Score 629.8; DB 1; Length 4309;

Best Local Similarity 57.1%; Pred. No. 6.1e-126;

Matches 1306; Conservative 0; Mismatches 912; Indels 69; Gaps 6;

QY 175 GCATTTCAGGAGGACTTTACCTACACCTGACCGCGGATGCTCAGTTCCTGGCTCCCGCC 234

Db 353 GCCTTTGACGAGCTGGATCTGGAGCTGCGGCCCGGACAGAGCTTTTGGCGCCGCGC 412

QY 235 TTCTCACTGAGCATCTGCGGCTCCCGCTCCAGGGGCTCACGGGGGCTC---TTTCAGAC 291

Db 413 TTCAGGCTCCAGACGCTGGGCGGCAATCGGCTCGGAGCGCGCTTCGGAACCGAC 472

QY 292 CTGCGACGCTCTTATCTGGGAGCTGAACGCGGACGCGGACTCGTTCTGCTCTG 351

Db 473 CTGGCGCACTGCTTACTCGGACCGCTGAATGGGATCCGACCTCGGCTCGGCGCTC 532

QY 352 AGCCTGTGCGGGGCTCGCGAGGCTTTGGTACCGAGGCGGAGTATCATAGC 411

Db 533 AGCCTGTGAGGGGCTGCGCGGCGCTTCTACCTGTGGGAGGCGTATTTCATCCAG 592

QY 412 CGGCTGCCAATGCTAGCGGC-CGGGCGGAGCGGACAGCCAGGCGGACACCTTCT 470

Db 593 CGGCTGCCCGCGGAGGCGCTCGCCACCGCGCCCGCGGAGGAGGAGCGCGGCA 652

QY 471 CCAGCGCGGGGTGTTCGGGCGGGCTTCCGGAGCCCACTCTCGTTCGGGGTGGC 530

Db 653 CCACCTACAGTTCACCTCTCGGCGGAATCGGCGGCGGACGTCGCGCGGCGG 712

QY 531 CTCGGCTGGACCCCGCTACCGGCTTACCGGCTTACAGCTCGCGGCGGCGG 590

Db 713 GTCGTGGACGAGGCGGCGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGG 772

QY 591 CTTTCGGGAGAGTCTGACCGGCGGCTGCTGGGCGC----- 627

Db 773 ACTGAGGCGGAGGAGGAGGCTCAGTGTGCGCGGAGGAGGAGGAGGAGGAGG 832

QY 628 -----GCCAGGCTTCGTCTATCCCGGCTAC 657

Db 833 GGACAGCCACAGGAGTGAAGCATGAAGAAAGAGCGATTGTGTCCAGTCAACGCTAT 892

QY 658 GTGGAGAGCGTGGTGGCGGAGTGAATGCTCAAGTTCACGCGGCTGACGCTGAA 717

Db 893 GTGGAAACCATGCTTGTGGACAGCAGTGCATGAGAGGAGGAGGAGGAGGAGG 952

QY 718 CATTATCTGCTGAGCTGCTGGCAACGCGGCGGAGCTTACCGGCTCCAGGATCTC 777

Db 953 CATTACCTTCTCAGCTTGTGTTTTCGGTGGCAGCCAGATTGTACAAACACCCAGCAATTCGT 1012

QY 778 AACCCCATCAACATCGTTGTGGTCAAGGTGCTGCTTCTTTAGAGATCGTGACTCCGGGCC 837

Db 1013 AATTCACTAGCTTGGTGGTGAAGATCTTGTGTCATCCACGATGAACAGGAGGCGCG 1072

QY 838 AAGTCCACCGCAATCGGCGCTGACGCTCGCAACTTCTGTGCTGCGGAGAGAGTGTG 897

Db 1073 GAAGTGACCTCCAAATGCTGCGCTCCTCCTCGGAACTTTTGGCACTGGCAGAGAGCAGC 1132

QY 898 AACAAAGTGAAGCAAGCAGCCCGAGTACTGGGACACTGCTCTCTTCACAGGAGC 957

Db 1133 AACCCACCCAGTGAACCGGATGACAGCACTATGACACAGCAATTTCTTTCACAGACAG 1192

QY 958 GACCTGTGTGGAGCCACCTGTGACACCTGGGATGCTGATGCTGGTACCATGTGT 1017

Db 1193 GACTGTGTGGTCCAGACATGATGCTTGGGATGCTGATGTTGAAGTGTGT 1252

QY 1018 GACCCCAAGAGAGTGTCTGCTCAATTGAGAGAGTGGCTTCCATCAGGCTTCACACT 1077

Db 1253 GATCCGAGCAGAAAGTGTCTCGTTCATAGAAGATGATGTTTACAAGCTGCCITTCACACA 1312

QY 1078 GCCCAGAGCTGGCGAGTGTTCACATGCCCATGCAATGTGAAGTCTGTAGGAG 1137

Db 1313 GCCCATGAATTAGCCAGCTGTTTAAATGCCACATGATGCAAGAGCTGTGACG 1372

QY 1138 GTTGTGGAGAGCTCCGAGCCACACATGATGTCGCCGACCTCATTCAGATCGACCT 1197

Db 1373 CTTAATGTGTGAACAGGATTCACATGATGCGCTCAATGCTTCCAACTTGGACAC 1432

QY 1198 GCAACCCCTGCTGAGCTGCTGAGTGTGCTGCTATCATCCGACTTCTTGGAGAGCGGAC 1257

Db 1433 AGCCAGCTTGTCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1492

QY 1258 GGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317

Db 1493 GGGGAATGTTGATGAGCAACCTCAGATCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552

QY 1318 GCGAGTCAACCTCAGCAGCAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377

Db 1553 ACCTGTAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1612

QY 1378 CTTTACATGCTATC---TGCACCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1434

Db 1613 CCGATGACGAGCAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1672

QY 1435 GTGTGCGAGCAGCAGCTTCCCTGCGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494

Db 1673 GTGTGCTAAACCAACACTTCCCTGCGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1732

QY 1495 TGCTTCAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1548

Db 1733 TGTATCAAGCGCAAGTGTGTGAACAAACCGACAGAAAGCATTTTATAGCTGCTTTCAT 1792

QY 1549 GGTTCCTGGCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1608

Db 1793 GGAAGTGGGATGTGGGCGCTTGGGAGACTTTCGAGAGCTGCTGCTGCTGCTGCTGCTGCT 1852

QY 1609 CAGCTGGCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1668

Db 1853 CAGTACATGAGGAGATGTGACACCCAGTCCCAAGAAATGAGGAGGAGTGTGCTGCTGCTGCT 1912

QY 1669 GGAGTGAAGGCAATACGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728

Db 1913 GGCAACAGTTCGCTACAGATCTGTAACCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1969

QY 1729 AAGAGCTTCGGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788

Db 1970 AAAACCTTTAGAGAGCAACATGTGAACACACACAGAGTTCCTTCAAGCTTCTTTGGG 2029

QY 1789 CTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1848

Db 2030 AGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2089

```
QY 1849 AAGCTCATCTGCGAGCCAAATGGCACTGGCTACTTCTATGTGTGGCAACCAAGGTGGTG 1908
    |||||
Db 2090 AAGCTCATCTGCGAGCCAAAGCAATGGCTACTTCTCGTTTTCAGCCCAAGGTGTGA 2149
    |||||
QY 1909 GACGGCACGCTGTCTCTCTGACTCCACTCCGCTCTGTCTCCAAAGCAAGTGCATCAAG 1968
    |||||
Db 2150 GATGGTACTCCATGTAGCCCAAGTCCCACTCTGTCTGTGCAAGGACAGTGTGTAAAA 2209
    |||||
QY 1969 GCTGGCTGTGATGGGAACCTGGGCTCCAAAGAGAGATTTCGACAAAGTGTGGGGTGTGGG 2028
    |||||
Db 2210 CTTGGTTGTGATCGCATCATAGACTCCAAAAGAAAGTTTGATAAATGTGGTGTGGCGG 2269
    |||||
QY 2029 GGAGACAATAAGAGCTGCAAGAGAGTGACTGGACTCTTCACCAAGCCCATGCATGGCTAC 2088
    |||||
Db 2270 GGAATGGATCTACTTGTAAAAAATATCAGGATCAGTTACTAGTGCAAAACCTGGATAT 2329
    |||||
QY 2089 AATTTCTGTGGTGGCATCCCGCGAGCGCCTCAAGCATCGACATCCGCGCAGCGGGGTAC 2148
    |||||
Db 2330 CATGATATCATCAAAATTCCACTGGAGGCCCAACATCGAAGTGAACACAGCGGAACCCAG 2389
    |||||
QY 2149 AAAGGGCTGATCGGGGATGACAACTACCTGGCTCTGAAGAACAGCCAAAGCAAGTACCTG 2208
    |||||
Db 2390 AGGGGATCCAGGAACAATGGCAGCTTCTTGCCATCAAAAGCTGCTGATGSCACATATTT 2449
    |||||
QY 2209 CTCACGGGCAATTTCTGTGTGTGCGGGTGGAGCGGGACCTGGTGTGAAGGCGAGTCTG 2268
    |||||
Db 2450 CTTAATGGTGACTACTTTTGCCCTTAGAGCAAGACATTATGTACAAGGTGTGTGTC 2509
    |||||
QY 2269 CTGCGGTACACGGGCACGGGCACAGCGTGTGAGAGCCTTCAGAGCTTCCCGGCCCATCCTG 2328
    |||||
Db 2510 TTGAGGTACACGGGCTCTCTCGGGCATTTGAAAGAAATTCGACGCTTAGCCCTCTCAAA 2569
    |||||
QY 2329 GAGCGGCTGACCGTGGAGGTCTCTCCGTGGGGAAGATGACACCGCCCGGTCGCGCTAC 2388
    |||||
Db 2570 GAGCCCTTGACCATCCAGGTCTTACTGTGGGCAATGCCCTTCGACCTAAATTAATATAC 2629
    |||||
QY 2389 TCCCTTCT 2395
    |||||
Db 2630 ACCTACT 2636
```

Search completed: May 16, 2003, 00:46:42
Job time : 1130 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:22:22 ; Search time 159 Seconds
(without alignments)
3852.179 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLGLILTLAFAGTAGGSE.....DQCNLHRKPQELDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main :

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US09_COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	5162	100.0	950	21	US-09-741-151-2	Sequence 2, Appli
2	5162	100.0	950	23	US-09-965-631-4	Sequence 4, Appli
3	5156	99.9	950	24	US-10-009-332-1	Sequence 1, Appli
4	5117	99.1	952	27	US-60-216-821-1	Sequence 4, Appli
5	5051	97.8	1132	27	US-60-212-656-497	Sequence 497, App
6	5051	97.8	1207	27	US-60-230-435-1155	Sequence 1155, App

7	5041	97.7	1252	27	US-60-242-679-1059	Sequence 1059, Ap
8	5038	97.6	1158	27	US-60-207-315-292	Sequence 292, App
9	4921	95.3	924	24	US-10-093-463-28	Sequence 28, Appli
10	4758	92.2	884	27	US-60-208-020-103	Sequence 103, App
11	4758	92.2	884	27	US-60-209-043-133	Sequence 133, App
12	4248.5	82.3	823	25	US-10-163-316-2	Sequence 2, Appli
13	4248.5	82.3	823	25	US-60-297-863-2	Sequence 2, Appli
14	3963.5	76.8	755	24	US-10-093-453-30	Sequence 30, Appli
15	3528	49.0	472	27	US-60-206-028-172	Sequence 172, App
16	2488	48.2	462	27	US-60-206-028-173	Sequence 173, App
17	2487.5	48.2	931	21	US-09-791-537-96880	Sequence 96880, A
18	2486.5	48.2	949	19	US-09-741-151-4	Sequence 4, Appli
19	2486.5	48.2	949	19	US-09-568-559-2	Sequence 2, Appli
20	2486.5	48.2	950	1	PCT-US00-14462A-2	Sequence 2, Appli
21	2486.5	48.2	950	1	PCT-US99-01313-2	Sequence 2, Appli
22	2486.5	48.2	950	16	US-09-235-810-2	Sequence 2, Appli
23	2486.5	48.2	950	17	US-09-373-658B-2	Sequence 2, Appli
24	2486.5	48.2	950	17	US-09-373-658B-2	Sequence 2, Appli
25	2486.5	48.2	950	17	US-09-373-658B-2	Sequence 2, Appli
26	2486.5	48.2	950	21	US-09-791-537-108327	Sequence 108327, Sequence 2, Appli
27	2486.5	48.2	950	23	US-09-989-687-2	Sequence 2, Appli
28	2486.5	48.2	950	27	US-08-171-503-2	Sequence 2, Appli
29	2486.5	48.2	967	12	US-08-845-496-2	Sequence 2, Appli
30	2486.5	48.2	967	17	US-09-388-279-2	Sequence 2, Appli
31	2486.5	48.2	967	17	US-09-388-280-2	Sequence 2, Appli
32	2486.5	48.2	967	21	US-09-791-537-41654	Sequence 41654, A
33	2486.5	48.2	967	24	US-10-057-084-2	Sequence 2, Appli
34	2486.5	48.2	967	25	US-10-105-929-2	Sequence 2, Appli
35	2486.5	48.2	967	25	US-10-115-286-2	Sequence 2, Appli
36	2486.5	48.2	968	1	PCT-US00-14462A-125	Sequence 125, App
37	2486.5	48.2	968	17	US-09-373-658B-125	Sequence 125, App
38	2486.5	48.2	968	17	US-09-373-658B-125	Sequence 125, App
39	2486.5	48.2	968	27	US-60-171-503-125	Sequence 125, App
40	2486.5	48.2	968	21	US-09-791-537-1136	Sequence 1136, App
41	2486.5	48.1	950	17	US-09-321-987B-4	Sequence 4, Appli
42	2485.5	48.1	951	16	US-09-264-585-17	Sequence 17, Appli
43	2485.5	48.1	968	25	US-10-163-316-7	Sequence 7, Appli
44	2485.5	48.1	968	27	US-60-297-863-7	Sequence 7, Appli
45	2482.5	48.1	967	21	US-09-791-537-49234	Sequence 49234, A

ALIGNMENTS

RESULT 1
US-09-741-151-2
; Sequence 2, Application US/09741151
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shiaoqing et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO01005
; CURRENT APPLICATION NUMBER: US/09/741,151
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Human
US-09-741-151-2

Query Match 100.0%; Score 5162; DB 21; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLGLILTLAFAGTAGGSEPERVVPPIRLDPDINGRYWRGPDSDGDLIFQIIAF 60
Db 1 MLLGLILTLAFAGTAGGSEPERVVPPIRLDPDINGRYWRGPDSDGDLIFQIIAF 60
QY 61 QEDFYHLTPDAQFLAPAFSTHGLVPLQSLTGSSDLRRCFYSGDVNAEPDSFAVSLC 120
|||||

```
Db 61 QEDFYLHLTPDAQFLAPAFSTHGLVPLQGLTGGSSDLRRCFYSGDVNAEPDSPAFAVSLC 120
QY 121 GGLRGAFYGAEEYVISPPLNAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180
Db 121 GGLRGAFYGAEEYVISPPLNAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180
QY 181 NPAILRALDYPKPRRAGFGEESRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 181 NPAILRALDYPKPRRAGFGEESRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWOKKLN 300
Db 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWOKKLN 300
QY 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSVSIEDDGLPSAFITTA 360
Db 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSVSIEDDGLPSAFITTA 360
QY 361 HELGHVFNMPHDNPKVCEEVFGKLRANHMSPITLQIDRANPWSACSAIITDFLDSHG 420
Db 361 HELGHVFNMPHDNPKVCEEVFGKLRANHMSPITLQIDRANPWSACSAIITDFLDSHG 420
QY 421 DCLLDQPSKPIISLPEDLPFGASYTISOCELAFGVGSKPCPYMOYCTKLWCTGKAKGQWVC 480
Db 421 DCLLDQPSKPIISLPEDLPFGASYTISOCELAFGVGSKPCPYMOYCTKLWCTGKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKRVDSGAKWDPYPCSRCTCGGVQLAR 540
Db 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKRVDSGAKWDPYPCSRCTCGGVQLAR 540
QY 541 ROCTNPTPANGKYGCEGVRYKSCNLEPCPSASGKSFREEQCEAFNGYNHSTNRLTILA 600
Db 541 ROCTNPTPANGKYGCEGVRYKSCNLEPCPSASGKSFREEQCEAFNGYNHSTNRLTILA 600
QY 601 VAWPKYSGVSPRDKCKLICRANGTYFYVLAPKVVYDGTLCSPDSTSVYCVGKCIKAGCD 660
Db 601 VAWPKYSGVSPRDKCKLICRANGTYFYVLAPKVVYDGTLCSPDSTSVYCVGKCIKAGCD 660
QY 661 GNLGSKKRFDCGCGGDNKSKCKVYGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Db 661 GNLGSKKRFDCGCGGDNKSKCKVYGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
QY 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLEPT 780
Db 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLEPT 780
QY 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHPPKDPGSPVLSLNSVLSNOVEQDDORPP 840
Db 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHPPKDPGSPVLSLNSVLSNOVEQDDORPP 840
QY 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGORTVPACDAHRPVETQACGECPCPTWELS 900
Db 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGORTVPACDAHRPVETQACGECPCPTWELS 900
QY 901 AWSPCSKSGRGFORSLKCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950
Db 901 AWSPCSKSGRGFORSLKCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950
```

RESULT 2

```
US-09-965-631-4
; Sequence 4, Application US/09965631
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-965-631-4
```

```
Query Match 100.0%; Score 5162; DB 23; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MLLGLITLAFAGTAGGSEPEVVPVIRLDPDINGRRYWRGPDGSGDGLIFQITAF 60
Db 1 MLLGLITLAFAGTAGGSEPEVVPVIRLDPDINGRRYWRGPDGSGDGLIFQITAF 60
QY 61 QEDFYLHLTPDAQFLAPAFSTHGLVPLQGLTGGSSDLRRCFYSGDVNAEPDSPAFAVSLC 120
Db 61 QEDFYLHLTPDAQFLAPAFSTHGLVPLQGLTGGSSDLRRCFYSGDVNAEPDSPAFAVSLC 120
QY 121 GGLRGAFYGAEEYVISPPLNAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180
Db 121 GGLRGAFYGAEEYVISPPLNAPAAQNSOGAHLQRRGVYJGPGSDPTSRGCVASGW 180
QY 181 NPAILRALDYPKPRRAGFGEESRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 181 NPAILRALDYPKPRRAGFGEESRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWOKKLN 300
Db 241 YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWOKKLN 300
QY 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSVSIEDDGLPSAFITTA 360
Db 301 KVS DKHPEYWDTAILFTRODLGCATCTDGLMADVGTMCDPKRSVSIEDDGLPSAFITTA 360
QY 361 HELGHVFNMPHDNPKVCEEVFGKLRANHMSPITLQIDRANPWSACSAIITDFLDSHG 420
Db 361 HELGHVFNMPHDNPKVCEEVFGKLRANHMSPITLQIDRANPWSACSAIITDFLDSHG 420
QY 421 DCLLDQPSKPIISLPEDLPFGASYTISOCELAFGVGSKPCPYMOYCTKLWCTGKAKGQWVC 480
Db 421 DCLLDQPSKPIISLPEDLPFGASYTISOCELAFGVGSKPCPYMOYCTKLWCTGKAKGQWVC 480
QY 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKRVDSGAKWDPYPCSRCTCGGVQLAR 540
Db 481 QTRHFPWADGTSCEGKLCILKACVERHNLNKRVDSGAKWDPYPCSRCTCGGVQLAR 540
QY 541 ROCTNPTPANGKYGCEGVRYKSCNLEPCPSASGKSFREEQCEAFNGYNHSTNRLTILA 600
Db 541 ROCTNPTPANGKYGCEGVRYKSCNLEPCPSASGKSFREEQCEAFNGYNHSTNRLTILA 600
QY 601 VAWPKYSGVSPRDKCKLICRANGTYFYVLAPKVVYDGTLCSPDSTSVYCVGKCIKAGCD 660
Db 601 VAWPKYSGVSPRDKCKLICRANGTYFYVLAPKVVYDGTLCSPDSTSVYCVGKCIKAGCD 660
QY 661 GNLGSKKRFDCGCGGDNKSKCKVYGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Db 661 GNLGSKKRFDCGCGGDNKSKCKVYGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
QY 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLEPT 780
Db 721 GDDNYLALKNSQGYLLNGHFVYSAVERDLVVKGSLLRYSGTGAVESLQASRPILPLEPT 780
QY 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHPPKDPGSPVLSLNSVLSNOVEQDDORPP 840
Db 781 VEVLVSGKMTPPRVRYSFYLPKPEPREDKSSHPPKDPGSPVLSLNSVLSNOVEQDDORPP 840
QY 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGORTVPACDAHRPVETQACGECPCPTWELS 900
Db 841 ARWVAGSWGPCASCSGSLQKRAVDCRGSGAGORTVPACDAHRPVETQACGECPCPTWELS 900
QY 901 AWSPCSKSGRGFORSLKCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950
Db 901 AWSPCSKSGRGFORSLKCVGHGGLLARDQCNLHRKPQELDFCVLRPC 950
```

RESULT 3			US-10-009-332-1		
; Sequence 1, Application US/10009332			; GENERAL INFORMATION:		
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.			; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY		
; FILE REFERENCE: Q67541			; CURRENT APPLICATION NUMBER: US/10/009,332		
; PRIOR FILING DATE: 2001-12-10			; PRIOR FILING DATE: 1999-11-11		
; PRIOR APPLICATION NUMBER: JPA 2000-144020			; PRIOR FILING DATE: 2000-05-16		
; NUMBER OF SEQ ID NOS: 35			; SOFTWARE: Patent version 3.1		
; SEQ ID NO 1			; LENGTH: 950		
; TYPE: PRT			; ORGANISM: Homo sapiens		
US-10-009-332-1			Query Match		
			Best Local Similarity 99.9%; Score 5156; DB 24; Length 950;		
			Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60			
DB	1	MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60			
QY	61	QEDFYHLTDAQFLAPAFSTEHLGVPLQGLTGGSSDLRCFTYSGDVNAEPDSEFAVSLC 120			
DB	61	QEDFYHLTDAQFLAPAFSTEHLGVPLQGLTGGSSDLRCFTYSGDVNAEPDSEFAVSLC 120			
QY	121	GGLRGAGYGAEVVISPPLNASAPAAORNSOGAHLIQRGVPGGSGDPTSRGCVASGW 180			
DB	121	GGLRGAGYGAEVVISPPLNASAPAAORNSOGAHLIQRGVPGGSGDPTSRGCVASGW 180			
QY	181	NPAILRALDPYKPRRAGGESRRSRGAKRFVSPRYVETLIVADESVVKFHGADLEH 240			
DB	181	NPAILRALDPYKPRRAGGESRRSRGAKRFVSPRYVETLIVADESVVKFHGADLEH 240			
QY	241	YLLTLLTAARLYRHPSTLNPINIVVYKLLLRDRDSGPXVTGNAATLNFCAWQKLN 300			
DB	241	YLLTLLTAARLYRHPSTLNPINIVVYKLLLRDRDSGPXVTGNAATLNFCAWQKLN 300			
QY	301	KVSDKHPEYDWTAILTRQDLGCGATTCDTLGMADVGTMCDPKRSVIEDDGLPSAFTTA 360			
DB	301	KVSDKHPEYDWTAILTRQDLGCGATTCDTLGMADVGTMCDPKRSVIEDDGLPSAFTTA 360			
QY	361	HELGHVFNMPHDNVKVEEVPFGKLRANHMSPPTLIQIDRANPWSACSAAIITDLDSCGH 420			
DB	361	HELGHVFNMPHDNVKVEEVPFGKLRANHMSPPTLIQIDRANPWSACSAAIITDLDSCGH 420			
QY	421	DCLLDQPSKPISLPDLPGASYTLSOQCELAFGVSGKPCPYMOYCTKLWCTGKAKGOMVC 480			
DB	421	DCLLDQPSKPISLPDLPGASYTLSOQCELAFGVSGKPCPYMOYCTKLWCTGKAKGOMVC 480			
QY	481	QTRHPFWADGTSCGSKGLCKLGACVERHNLNKRVDGSKWAKWDPYPCSRITCGGGVQLAR 540			
DB	481	QTRHPFWADGTSCGSKGLCKLGACVERHNLNKRVDGSKWAKWDPYPCSRITCGGGVQLAR 540			
QY	541	ROCTNPTPANGKYCEGVVRYRSCNLEPCPSSASGKSFIEQCEAFNGYNHSTNRLTLA 600			
DB	541	ROCTNPTPANGKYCEGVVRYRSCNLEPCPSSASGKSFIEQCEAFNGYNHSTNRLTLA 600			
QY	601	VAMVPKYSVSPDRCKKLCIRANGTGYFYVLAPKVVDGTCLSPDSTSVCVGKCIKAGCD 660			
DB	601	VAMVPKYSVSPDRCKKLCIRANGTGYFYVLAPKVVDGTCLSPDSTSVCVGKCIKAGCD 660			
QY	661	GNLGSKRFRDKGCGVCGDNKSKKVTGLTKPMHGYNFVVAIPAGASSIDIRQRYKGLI 720			
US-09-965-631-4			Query Match		
			Best Local Similarity 99.1%; Score 5117; DB 27; Length 954;		
			Matches 947; Conservative 0; Mismatches 3; Indels 2; Gaps 2;		
QY	1	MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60			
DB	1	MLLGLTLTAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60			
QY	61	QEDFYHLTDAQFLAPAFSTEHLGVPLQGLTGGSSDLRCFTYSGDVNAEPDSEFAVSLC 120			
DB	61	QEDFYHLTDAQFLAPAFSTEHLGVPLQGLTGGSSDLRCFTYSGDVNAEPDSEFAVSLC 120			
QY	121	GGLRGAGYGAEVVISPPLNASAPAAORNSOGAHLIQRGVPGGSGDPTSRGCVASGW 180			

```

121 GGLRGAFGYRGAAYVISPLPNASAPAAQNRSGQAHLLQRRGVPGGSGDPTSRGCVASGW 180
181 NPAILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
181 NPAILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
241 YLLTLATAARLYRHPSTILNINIVVKKVLLRRDSDGPKVTGNAALTLENCFAWQKKLN 300
241 YLLTLATAARLYRHPSTILNINIVVKKVLLRRDSDGPKVTGNAALTLENCFAWQKKLN 300
301 KYSDKHPEYWDTAIFTTQDLGCGATTCTDLGADVGTMCDDPKRSCSVIEDDGLPSAFTTA 360
301 KYSDKHPEYWDTAIFTTQDLGCGATTCTDLGADVGTMCDDPKRSCSVIEDDGLPSAFTTA 360
361 HELGHVFNPHNDNVKVEEFGKLRANHMSPPTLIQIDRANPWSACSAAIITDFLDSGGH 420
361 HELGHVFNPHNDNVKVEEFGKLRANHMSPPTLIQIDRANPWSACSAAIITDFLDSGGH 420
421 DCLLDQPSKPISLPDLPGASVTLISQOCELAFGVGSKPCPYMOYCTKLWCTGKAKGQWVC 480
421 DCLLDQPSKPISLPDLPGASVTLISQOCELAFGVGSKPCPYMOYCTKLWCTGKAKGQWVC 480
481 QTRHFPWADGTCGSGKGLKLGACVERHNLNKRVDSGAKWMDPYGPCSRTCGGCVOLAR 540
481 QTRHFPWADGTCGSGKGLKLGACVERHNLNKRVDSGAKWMDPYGPCSRTCGGCVOLAR 540
541 ROCTNPTPANGKYGCEGVYRVYRSCNLEPCPSSASGSKSFREOCEAFNGYHNSTNRLTLA 600
541 ROCTNPTPANGKYGCEGVYRVYRSCNLEPCPSSASGSKSFREOCEAFNGYHNSTNRLTLA 600
601 VAWPKYSGVSPRDKCLICRANGTGYFYVLAPK-VVDGTLCSPDSTSVCVQKCIKAGC 659
601 VAWPKYSGVSPRDKCLICRANGTGYFYVLAPK-VVDGTLCSPDSTSVCVQKCIKAGC 660
660 DGNLGSKKRFPDKCGVCGGDNKSKKVTGLTKPMHGNFVVAIPAGASSIDIRORGKYL 719
661 DGNLGSKKRFPDKCGVCGGDNKSKKVTGLTKPMHGNFVVAIPAGASSIDIRORGKYL 720
720 IGDNDYALKNSQGYLLNGHFVSAVERDLVVKGSLRLRYSGTGAVESLQASRPILLEPL 779
721 IGDNDYALKNSQGYLLNGHFVSAVERDLVVKGSLRLRYSGTGAVESLQASRPILLEPL 780
780 TVEVLSVGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDR-GPSVLHNSVLNSNOVEQDDR 838
781 TVEVLSVGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDR-GPSVLHNSVLNSNOVEQDDR 840
839 PPARWAGSWGPCASGSGGLQKRAVDCRGAGORTVPACDAHRPVETACGCEPCTWE 898
841 PPARWAGSWGPCASGSGGLQKRAVDCRGAGORTVPACDAHRPVETACGCEPCTWE 900
899 LSAWSPCKSCGRGFORSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950
901 LSAWSPCKSCGRGFORSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 952
```

```

RESULT 5
US-60-212-656-497
; Sequence 497, Application US/6012656
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000673
; CURRENT APPLICATION NUMBER: US/60/212,656
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 497
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: HUMAN
```

```

US-60-212-656-497
Query Match 97.8%; Score 5051; DB 27; Length 1132;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 182; Gaps 2;

QY 1 MLLGILTLAFAGTAGGSEPERVVPIRLDPDINRRYRWRCPEDSGGGLIFQITAF 60
DB 1 MLLGILTLAFAGTAGGSEPERVVPIRLDPDINRRYRWRCPEDSGGGLIFQITAF 60
QY 61 QDEDFYHLTPDAQFLAPAFSTEHLGVLPQLGTGSSDLRRCFCYSGDVNAEPDFAAVSLC 120
DB 61 QDEDFYHLTPDAQFLAPAFSTEHLGVLPQLGTGSSDLRRCFCYSGDVNAEPDFAAVSLC 120
QY 121 GGLRGAFGYRGAAYVISPLPNASAPAAQNRSGQAHLLQRRGVPGGSGDPTSRGCVASGW 180
DB 121 GGLRGAFGYRGAAYVISPLPNASAPAAQNRSGQAHLLQRRGVPGGSGDPTSRGCVASGW 180
QY 181 NPAILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
DB 181 NPAILRALDPYKPRRAGFGESESRSSRRSRGAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLATAARLYRHPSTILNINIVVKKVLLRRDSDGPKVTGNAALTLENCFAWQKKLN 300
DB 241 YLLTLATAARLYRHPSTILNINIVVKKVLLRRDSDGPKVTGNAALTLENCFAWQKKLN 300
QY 301 KYSDKHPEYWDTAIFTTQDLGCGATTCTDLGADVGTMCDDPKRSCSVIEDDGLPSAFTTA 317
DB 301 KYSDKHPEYWDTAIFTTQDLGCGATTCTDLGADVGTMCDDPKRSCSVIEDDGLPSAFTTA 360
QY 318 ----- 317
DB 318 ----- 360
QY 361 IRRRGAGTQEAAPQLQVSKTDRRRREDGVGEGNGPPLLGRLLSLHLLRSPFKGCR 420
QY 318 -----RGEL 321
DB 421 TEGDPRCSNPKLFYKKGQLENNSPGGLRFLAKMIPGSSGKVAATAKNSQLLIVRCDL 480
QY 322 CGATTCTDGLMADVGTMCDDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHNDNVKVEEVF 381
DB 481 CGATTCTDGLMADVGTMCDDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHNDNVKVEEVF 540
QY 382 GKLRANHMSPPTLIQIDRANPWSACSAAIITDFLDSGGDLDDPSKPISLPDLPGAS 441
DB 541 GKLRANHMSPPTLIQIDRANPWSACSAAIITDFLDSGGDLDDPSKPISLPDLPGAS 600
QY 442 YTLSSQCELAFGVSKPCPYMOYCTKLWCTGKAKGQWVCQTRHFPWADGTCGEGKLCIK 501
DB 601 YTLSSQCELAFGVSKPCPYMOYCTKLWCTGKAKGQWVCQTRHFPWADGTCGEGKLCIK 660
QY 502 GACVERHNLNKHR-----VDGSAKWWDYPGCSRTCGGCVQL 538
DB 661 GACVERHNLNKHRALTDIISPKQLLLLPNGLHTTKYDGSNAKWDPYGPCSRTCGGCVQL 720
QY 539 ARROCTNPTPANGKYGCEGVYRVYRSCNLEPCPSSASGSKSFREOCEAFNGYHNSTNFLT 598
DB 721 ARROCTNPTPANGKYGCEGVYRVYRSCNLEPCPSSASGSKSFREOCEAFNGYHNSTNFLT 780
QY 599 LAVAWPKYSGVSPRDKCLICRANGTGYFYVLAPKVVDTGLCSPDSTSVCVQKCIKAG 658
DB 781 LAVAWPKYSGVSPRDKCLICRANGTGYFYVLAPKVVDTGLCSPDSTSVCVQKCIKAG 840
QY 659 CDGNLGSKKRFPDKCGVCGGDNKSKKVTGLTKPMHGNFVVAIPAGASSIDIRORGKYL 718
DB 841 CDGNLGSKKRFPDKCGVCGGDNKSKKVTGLTKPMHGNFVVAIPAGASSIDIRORGKYL 900
QY 719 LIGDNDYALKNSQGYLLNGHFVSAVERDLVVKGSLRLRYSGTGAVESLQASRPILLEP 778
DB 901 LIGDNDYALKNSQGYLLNGHFVSAVERDLVVKGSLRLRYSGTGAVESLQASRPILLEP 960
QY 779 LTVEVLSVGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDRGPSVLHNSVLNSNOVEQDDR 838
DB 961 LTVEVLSVGKMTPPRYRYFYFLPKPEPREDKSSHPKDPDRGPSVLHNSVLNSNOVEQDDR 1020
```

Qy 839 PPARWAGSWGPGSCSGGLQKRAVDCRGSGAGQRTVPACDAHRPVETACGECPTWE 998
Db 1021 PPARWAGSWGPGSCSGGLQKRAVDCRGSGAGQRTVPACDAHRPVETACGECPTWE 1080
Qy 899 LSWSKSCSGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 950
Db 1081 LSWSKSCSGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 1132

RESULT 6
US-60-230-435-1155
; Sequence 1155, Application US/60230435;
; GENERAL INFORMATION:
; APPLICANT: Reasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE OF INVENTION: US/60230435
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1155

Query Match 97.8%; Score 5051; DB 27; Length 1207;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 182; Gaps 2;

Qy 1 MLLGILTLTAPAGTAGGSEPERVVPRLDPDINGRYYWRGPDGSGGLIFQITAF 60
Db 76 MLLGILTLTAPAGTAGGSEPERVVPRLDPDINGRYYWRGPDGSGGLIFQITAF 135
Qy 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSDLRRCFYSGDVNAEFDGFAAVSLC 120
Db 136 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSDLRRCFYSGDVNAEFDGFAAVSLC 195
Qy 121 GGLRGAGYGAEEVYISPLPNASAPAAQNSOGAHLRORRGPVGGSDTISRGGVAGW 180
Db 196 GGLRGAGYGAEEVYISPLPNASAPAAQNSOGAHLRORRGPVGGSDTISRGGVAGW 255
Qy 181 NPALLRALDPYKRRAGFGESESRSSRRSGRAKRFVSIPIRYVETLVVADESVMKPHGADLEH 240
Db 256 NPALLRALDPYKRRAGFGESESRSSRRSGRAKRFVSIPIRYVETLVVADESVMKPHGADLEH 315
Qy 241 YLLTLATAARLYRHPHSILNPINIVVVKVLLLRDRSGPKVTGNAALTLRNFCAWQKLN 300
Db 316 YLLTLATAARLYRHPHSILNPINIVVVKVLLLRDRSGPKVTGNAALTLRNFCAWQKLN 375
Qy 301 KVSDEHPYWDITALET----- 317
Db 376 KVSDEHPYWDITALETTRQPGESPLPSRAERRLTICWAGEPALSGVGLARPISASW 435
Qy 318 ----- 317
Db 436 IRRRGAGTQEAPELQFVSKTDRRRRRDEGVGEGNGPPLLYGRLSLGLHLRSPFKGCR 495
Qy 318 -----RQDL 321
Db 496 TEGDPRGSGNPKLFYKKGQLENNSPGGLRFLAKMIPGSSGKVAATAKNSQLLGIIVRQDL 555
Qy 322 CGATTCDTLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHONVKVCEVF 381
Db 556 CGATTCDTLGMADVGMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHONVKVCEVF 615
Qy 382 GKLRANHMSPTLIQIDRANPWSACSAIITDFLDGSGHGDCLLDQPSKPSLPEDLFGAS 441
Db 616 GKLRANHMSPTLIQIDRANPWSACSAIITDFLDGSGHGDCLLDQPSKPSLPEDLFGAS 675

Qy 442 YTLSSQCELAFAVGSKPCPYMOYCTKLACTGKAKGMVYCTRHFPWADGTSCEGKLCIK 501
Db 676 YTLSSQCELAFAVGSKPCPYMOYCTKLACTGKAKGMVYCTRHFPWADGTSCEGKLCIK 735
Qy 502 GACVERHLNKHRR-----VDGSAKWDPYGPGSRTRTCGGVQL 538
Db 736 GACVERHLNKHRRALDIIISPKQLLRLPGLLTHTKVDGSAKWDPYGPGSRTRTCGGVQL 795
Qy 539 ARROCTNPTPANGKYCEGVYKVRKSCNLEPCSPSSASGSKSFRECEAFNHNHSTNRLT 598
Db 796 ARROCTNPTPANGKYCEGVYKVRKSCNLEPCSPSSASGSKSFRECEAFNHNHSTNRLT 855
Qy 599 LAVAWPKYSGVSPRDKCKLICRANGTGYFYVLAPVKVVDGTLCSPDSTSVCVQKCIKAG 658
Db 856 LAVAWPKYSGVSPRDKCKLICRANGTGYFYVLAPVKVVDGTLCSPDSTSVCVQKCIKAG 915
Qy 659 CDGNLGSKKRFRDKCGVCGGDKNSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGVKG 718
Db 916 CDGNLGSKKRFRDKCGVCGGDKNSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGVKG 975
Qy 719 LIGDDNYLALKNSQGYLYLNGHFVSAVERDLVWKGSLRLRYSGTAVESLQASRPILPEP 778
Db 976 LIGDDNYLALKNSQGYLYLNGHFVSAVERDLVWKGSLRLRYSGTAVESLQASRPILPEP 1035
Qy 779 LTVEVLVSGKMTTPRVRYSPYLPKPEPRDKSSHPRGPGSVLHNSVLSLNOVQPDQR 838
Db 1036 LTVEVLVSGKMTTPRVRYSPYLPKPEPRDKSSHPRGPGSVLHNSVLSLNOVQPDQR 1095
Qy 839 PPARWAGSWGPGSCSGGLQKRAVDCRGSGAGQRTVPACDAHRPVETACGECPTWE 898
Db 1096 PPARWAGSWGPGSCSGGLQKRAVDCRGSGAGQRTVPACDAHRPVETACGECPTWE 1155
Qy 899 LSWSKSCSGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 950
Db 1156 LSWSKSCSGRGFORRSKLCVGHGRLARDQCNLHRRKPOELDFCVLRPC 1207

RESULT 7
US-60-242-679-1059
; Sequence 1059, Application US/60242679
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steven Istvan
; APPLICANT: Spier, Eugene
; APPLICANT: Greenberg, Simon
; APPLICANT: Brandenberger, Ralph
; APPLICANT: Wang, Yu
; APPLICANT: Dubman, Alex
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE OF INVENTION: US/60/242,679
; CURRENT APPLICATION NUMBER: US/60/242,679
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 2265
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1059
; LENGTH: 1252
; TYPE: PRT
; ORGANISM: HUMAN
US-60-242-679-1059

Query Match 97.7%; Score 5041; DB 27; Length 1232;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 202; Gaps 2;

Qy 1 MLLGILTLTAPAGTAGGSEPERVVPRLDPDINGRYYWRGPDGSDGGLIFQITAF 60
Db 101 MLLGILTLTAPAGTAGGSEPERVVPRLDPDINGRYYWRGPDGSDGGLIFQITAF 160
Qy 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSDLRRCFYSGDVNAEFDGFAAVSLC 120
Db 161 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSDLRRCFYSGDVNAEFDGFAAVSLC 220

```
QY 121 GGLRGAGYGAETVLSPLPNASAPAAQNSOGAHLQRRGVPGGPSDPTSCRGVASGW 180
Db 221 GGLRGAGYGAETVLSPLPNASAPAAQNSOGAHLQRRGVPGGPSDPTSCRGVASGW 280
QY 181 NPAILRALDYPKRRAGFGESESRSSRAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 281 NPAILRALDYPKRRAGFGESESRSSRAKRFVSIPIRYVETLVVADESVMKFGADLEH 340
QY 241 YLLTLATAARLYRHPSILNPINIVVVKLLRDSDGPKVTGNAALTLENFCAWOKKLN 300
Db 341 YLLTLATAARLYRHPSILNPINIVVVKLLRDSDGPKVTGNAALTLENFCAWOKKLN 400
QY 301 KVS DKHPEYWDTAILFTQ----- 319
Db 401 KVS DKHPEYWDTAILFTQ----- 460
QY 320 ----- 319
Db 461 IRRRGAGTQVAAPELQFVSKTDRRRRDGVGEGNGGPPILLYGRSLSLGLHLLRSPFKGCR 520
QY 320 -----DLGATTCDTLGMADVGTMCOPKRS 344
Db 521 TEGLSPLSVNRHLELQGEAFHSWGDHQRQOEVDMLCGATTCDTLGMADVGTMCOPKRS 580
QY 345 CSVIEDGLPSAFTTAHELGHVFNPHDNVNYKCEVEFGKLRANHMMSPTLIQIDRANPWS 404
Db 581 CSVIEDGLPSAFTTAHELGHVFNPHDNVNYKCEVEFGKLRANHMMSPTLIQIDRANPWS 640
QY 405 ACSAAITDPLDSHGDCILLDQSPKIPSLPDLPGASTYLSQOCELAFAFGVSGKPCPYMQY 464
Db 641 ACSAAITDPLDSHGDCILLDQSPKIPSLPDLPGASTYLSQOCELAFAFGVSGKPCPYMQY 700
QY 465 CTKLWCTGKAGQVMQVOTRHFHWADGTSCGEGKCLKACVERHNLNKH----- 514
Db 701 CTKLWCTGKAGQVMQVOTRHFHWADGTSCGEGKCLKACVERHNLNKHRLSLMTFLLIS 760
QY 515 -----VDGS 518
Db 761 RRGKPTGSKELYKLIAVNSIPFLPSAAHSQAOLDIISPKQLLLRLPNGLHTTKVDGS 820
QY 519 WAKWDYPGCSRTCCGGVQLARRQCTNPTPANGGKYCEGVVYKRSCLNLEPCPSASGKS 578
Db 821 WAKWDYPGCSRTCCGGVQLARRQCTNPTPANGGKYCEGVVYKRSCLNLEPCPSASGKS 880
QY 579 FREEQCEAFNGYNSHSTNRLTLAVAWPKYSGVSPDKCKLICRANGTGYFVLAPKVVDG 638
Db 881 FREEQCEAFNGYNSHSTNRLTLAVAWPKYSGVSPDKCKLICRANGTGYFVLAPKVVDG 940
QY 639 TLCSPDSTSVQVQCKIKAGCDGNLGSKKRDKCGVCGDKNKCKKVTGLTKPMHGYNF 698
Db 941 TLCSPDSTSVQVQCKIKAGCDGNLGSKKRDKCGVCGDKNKCKKVTGLTKPMHGYNF 1000
QY 699 VVAIPAGASSIDIRQRYKGLIGDDNYLALNKSQKYLNGHFVVSVAVERDLVYKGSLLR 758
Db 1001 VVAIPAGASSIDIRQRYKGLIGDDNYLALNKSQKYLNGHFVVSVAVERDLVYKGSLLR 1060
QY 759 YSGTGTAVESLQASRPILSPLVEVLSVCKMTPPRVRYSYFLPKREPREDKSSHPKOPRGP 818
Db 1061 YSGTGTAVESLQASRPILSPLVEVLSVCKMTPPRVRYSYFLPKREPREDKSSHPKOPRGP 1120
QY 819 SVLHNSVLISNVOEPPDRPARVAGSWGPCASCSGSLQKRAVDCRCAGQRTVPAC 878
Db 1121 SVLHNSVLISNVOEPPDRPARVAGSWGPCASCSGSLQKRAVDCRCAGQRTVPAC 1180
QY 879 DAHRPVETOACGEPCTWELSAWSPCKSKSGRGFQRRSLKCVCHGGRLLARQCNLHRK 938
Db 1181 DAHRPVETOACGEPCTWELSAWSPCKSKSGRGFQRRSLKCVCHGGRLLARQCNLHRK 1240
QY 939 PQELDFCVLRPC 950
Db 1241 PQELDFCVLRPC 1252
```

```
RESULT 8
US-60-207-315-292
; Sequence 292, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 292
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-292

Query Match 97.6%; Score 5038; DB 27; Length 1158;
Best Local similarity 82.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 208; Gaps 2;

QY 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGQGLIFOITAF 60
Db 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGQGLIFOITAF 60
QY 61 QEDFYHLTPDQAQFLAPAFSTHGLGVPIQGLTGGSSDLRCFCYSGDVNAEPDSFAAVSLC 120
Db 61 QEDFYHLTPDQAQFLAPAFSTHGLGVPIQGLTGGSSDLRCFCYSGDVNAEPDSFAAVSLC 120
QY 121 GGLRGAGYGAETVLSPLPNASAPAAQNSOGAHLQRRGVPGGPSDPTSCRGVASGW 180
Db 121 GGLRGAGYGAETVLSPLPNASAPAAQNSOGAHLQRRGVPGGPSDPTSCRGVASGW 180
QY 181 NPAILRALDYPKRRAGFGESESRSSRAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
Db 181 NPAILRALDYPKRRAGFGESESRSSRAKRFVSIPIRYVETLVVADESVMKFGADLEH 240
QY 241 YLLTLATAARLYRHPSILNPINIVVVKLLRDSDGPKVTGNAALTLENFCAWOKKLN 300
Db 241 YLLTLATAARLYRHPSILNPINIVVVKLLRDSDGPKVTGNAALTLENFCAWOKKLN 300
QY 301 KVS DKHPEYWDTAILFT----- 317
Db 301 KVS DKHPEYWDTAILFT----- 317
QY 318 ----- 317
Db 361 TCITWAGGEPALSEGVLARPIASWIRRRGAGTQVAAPELQFVSKTDRRREDGVGEG 420
QY 318 ----- 317
Db 421 GNGPPLLYGLSLILGHLHLLRSPKCGTSDPCGSKNPKLFYKKGQLENNSPGSLRFLAKM 480
QY 318 -----RDLCGATTCDTLGMADVGTMCOPKRSVIEDDGLS 355
Db 481 IPGSSGKVAATAKNSQLLGIYRQDLGATTCDTLGMADVGTMCOPKRSVIEDDGLS 540
QY 356 AFTTAHELGHVFNPHDNVNYKCEVEFGKLRANHMMSPTLIQIDRANPWSACSAAITDEL 415
Db 541 AFTTAHELGHVFNPHDNVNYKCEVEFGKLRANHMMSPTLIQIDRANPWSACSAAITDEL 600
QY 416 DSHGDCILLDQSPKIPSLPDLPGASTYLSQOCELAFAFGVSGKPCPYMQYCTKLWCTGKAK 475
Db 601 DSHGDCILLDQSPKIPSLPDLPGASTYLSQOCELAFAFGVSGKPCPYMQYCTKLWCTGKAK 660
QY 476 GQWVCOTRHFHWADGTSCGEGKCLKACVERHNLNKH----- 514
Db 661 GQWVCOTRHFHWADGTSCGEGKCLKACVERHNLNKHRLTIIISPKQLLLRLPNGLHT 720
QY 515 --VDGSWAKWDYPGCSRTCCGGVQLARRQCTNPTPANGGKYCEGVVYKRSCLNLEPCPS 572
Db 515 -----VDGSWAKWDYPGCSRTCCGGVQLARRQCTNPTPANGGKYCEGVVYKRSCLNLEPCPS 572
```


Db	721	TKVDGSAKWDPYGPCSRTCCGGVOLARQCTNTPPANGKYCSGVVRVKYRSCNLEPCPS	780
Qy	573	SASGKSFREQCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCKLICRANGTGYFYVLA	632
Db	781	SASGKSFREQCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCKLICRANGTGYFYVLA	840
Qy	633	PKYVDGTLCSPDSTSVCVQCGKCIKAGCDGNLGSKKRPDKGVCGGDNKSCKKVTGLETKP	692
Db	841	PKYVDGTLCSPDSTSVCVQCGKCIKAGCDGNLGSKKRPDKGVCGGDNKSCKKVTGLETKP	900
Qy	693	MHGYNFWALPAGASSIDIRQRYKGLIGDDNYLALKNSJCKYLLNGHFVYSAVERDLVV	752
Db	901	MHGYNFWALPAGASSIDIRQRYKGLIGDDNYLALKNSQKYLLNGHFVYSAVERDLVV	960
Qy	753	KGSLRYSGTGTAVESLQASRPLEPLTVEVLSVSGKMTPPRVRYSFYLPKXPREDKSSHP	812
Db	961	KGSLRYSGTGTAVESLQASRPLEPLTVEVLSVSGKMTPPRVRYSFYLPKXPREDKSSHP	1020
Qy	813	KDPRGSPVLHNSVLSLSNOVEQDDPPPARWYAGSWGSPCSASCGLQKXAVDCRGSAQO	872
Db	1021	KDPRGSPVLHNSVLSLSNOVEQDDPPPARWYAGSWGSPCSASCGLQKXAVDCRGSAQO	1080
Qy	873	RTVPACDAAHRPVETQACGPECTWELSAWSPCSKSCGRGQRSLKCVGHGGILLARDQ	932
Db	1081	RTVPACDAAHRPVETQACGPECTWELSAWSPCSKSCGRGQRSLKCVGHGGILLARDQ	1140
Qy	933	CNLRKPKQELDFCVLRPC	950
Db	1141	CNLRKPKQELDFCVLRPC	1158

RESULT 9

US-10-093-463-28

: Sequence 28, Application US/10093463

: GENERAL INFORMATION:

APPLICANT: Padigaru, Muralidhara	APPLICANT: Shenoy, Suresh
APPLICANT: Kekuda, Ramesh	APPLICANT: Gusev, Vladimir
APPLICANT: Pochart, Pascal	APPLICANT: Zhong, Mei
APPLICANT: Rastelli, Luca	APPLICANT: Mezes, Peter
APPLICANT: Smithson, Glennnda	APPLICANT: Guo, Xiaojia
APPLICANT: Gerlach, Valerie	APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc	APPLICANT: Li, Li
APPLICANT: Zerhusen, Bryan	APPLICANT: Tcheringev, Velizar
APPLICANT: Gangolli, Esha	APPLICANT: Vernet, Corinne
APPLICANT: Pena, Carol	APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong	APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda	APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward	APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David	APPLICANT: Patturajan, Meera
APPLICANT: Miller, Charles	APPLICANT: Taupier, Raymond J.
TITLE OF INVENTION: Novel Antibio	TITLE OF INVENTION: Encoding Th
FILE REFERENCE: 21A02-290A (Cura	CURRENT APPLICATION NUMBER: US/1/1
CURRENT FILING DATE: 2002-06-24	PRIOR APPLICATION NUMBER: 60/2838
PRIOR FILING DATE: 2001-04-14	PRIOR APPLICATION NUMBER: 60/338

Query Match

Query Match	95.3%	score 4921,
Best Local similarity	95.9%	Pred No. 0:

BEST LOCAL SIMILARITY 33.3%, FREQ. NO. 0,
Matches 914. Conservative 2: Mismatches 5: Indels 32: Gaps 3:

1 MLIIGTITIAFAGRTAGGSEPEREVVPTRI.DPDPINGRRYYWRGPEDSGDOGLIFOTAF 60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

61 CEDEVYHI.TPDPAOFI.APAFSTEHI.GVPI.OGI.TGGSSDI.RRCFYSGDVNAEPDSEFAAVSLC 120

QY I QEDF I EIBHIT ? DAQT EAT RI OI CENUE I OCEHNE I CEHNE I CEHNE I

D0 6T QEDFI LHLI FPAQF LAPAFSIEHLOVFLQUGLNGSSDLEKACTGSGDVANHEDEZAHNVDSIZ

121 GGI.RGA.FGYR.GAE.FY.VIS.PL.NAS.APA.AORN.SOG.AH.LLOR.RGV.PG.PSGD.PTS.F.CGVAS.GW 18

127 CCTPCAEGYPCAFVVTSPIDNASAPAAORNCSGAHLIORRGVPGGPGSDPSTRCGVASGW 180

00 170600Z NOV 66 A TROPOIC THUNDERSTORM

[illegible]

Db 181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESVMKFHGDLEH 240

QY 241 YI.T.T.I.L.A.TAARLYRHPSILNPINIVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300

22

THE UNIVERSITY OF CHICAGO PRESS

QV 301 KVS DKHPEYWD TAILFTRQDL CGATTCD TLGMADVGTMCDPKRSCSVIEDDGLPSAFTA 360

QY 421 DCLLDQPSKPISLPEDLPAGSYTILSQOCELAFAVGSGSKPCPYMQYCTKLWCTGKRAQGMVC 480
Db 421 DCLLDQPSKPISLPEDLPAGSYTILSQOCELAFAVGSGSKPCPYMQYCTKLWCTGKRAQGMVC 480
QY 481 QTRHPPNADGTSCEGKLCILKAGCVERHNLNKH---RVDGSWAKWDPYGPCSRTCCGGVQ 537
Db 481 QTRHPPNADGTSCEGKLCILKAGCVERHNLNKHSSQVDSWAKWDPYGFCSRTCCGGVQ 540
QY 538 LARQCTNPTPANGGKYCEGVVRKYRSCNLEPCPSSASGSKFEECEAFNGYNHSTNRL 597
Db 541 LARQCTNPTPANGGKYCEGVVRKYRSCNLEPCPS--SGKSFEEQCEAFNGYNHSTNRL 598
QY 598 TLAVAWPKYSGVSPRDKCLICRANGTGYYVLAKPVVDGTLCSPDSTSVCVQGCICA 657
Db 599 TLAVAWPKYSGVSPRDKCLICRANGTGYYVLAKPVVDGTLCSPDSTSVCVQGCICA 658
QY 658 GCDNLGSKKRFDKCGVCGDNKCKVTGLFTKPMHGYNFVVAIPAGASSIDIORGYK 717
Db 659 GCDNLGSKKRFDKCGVCGDNKCKVTGLLSPARHGYNFVVAIPAGASSIDIORGYK 718
QY 718 GLIGDDNYLALKNSQGYLLNGHFVSAVERDLVWKGSLRLRYSGTGTAVESLQASRPIL 777
Db 719 GLIGDDNYLALKNSQGYLLNGHFVSAVERDLVWKGSLRLRYSGTGTAVESLQASRPIL 778
QY 778 PLTVEVLSVCKMTPPRVRYSYFLPKPEPRDKSSHPRDKPRGSPVLHNSVLSLSNOVEOPDD 837
Db 779 PLTVEVLSVCKMTPPRVRYSYFLPKPEPRDKSSH-----812
QY 838 RPARWVAGSWGPCASCSGSLGKRAVDRCGSAGQRTVPACDAHRPVTQACGEPCTW 897
Db 813 -PRAWVAGSWGPCASCSGSLGKRAVDWEGSAGQRTVPACDAHRPVTQACGEPCTW 871
QY 898 ELSAWSPCSKSGRGQRRLSKCVGHGGRLLRDQCNLHRKPOELDFCVLRPC 950
Db 872 ELSAWSPCSKSGRGQRRLSKCVGHGGRLLRDQCNLHRKPOELDFCVLRPC 924

RESULT 10

US-60-208-020-103

; Sequence 103, Application US/60208020

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL000630

; CURRENT APPLICATION NUMBER: US/60/208,020

; NUMBER OF SEQ ID NOS: 178

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 103

; LENGTH: 884

; TYPE: PRT

; ORGANISM: HUMAN

US-60-208-020-103

Query Match 92.2%; Score 4758; DB 27; Length 884;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 878; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 44 GPDSGDOGLIFQITAFQEDFYHLHTPDQAFLAPAFSTHGLGVPLQGLTGSSDLRRCFY 103
Db 1 GPDSGDOGLIFQITAFQEDFYHLHTPDQAFLAPAFSTHGLGVPLQGLTGSSDLRRCFY 60
QY 104 SGDVNAEPDSFAVSLCGGLRGAFYRGAEYVISPPLNASPAQAQNSQGAHLQRRGVP 163
Db 61 SGDVNAEPDSFAVSLCGGLRGAFYRGAEYVISPPLNASPAQAQNSQGAHLQRRGVP 120
QY 164 GPGSDPTSRGCVASGWNPAIRALDPYKPRAGGSESRSSRRSGRAKRFVSIPIRYVETL 223
Db 121 GPGSDPTSRGCVASGWNPAIRALDPYKPRAGGSESRSSRRSGRAKRFVSIPIRYVETL 180
QY 224 VVADESMWKFHGCADLEHYLLTLLATAARLYRHPSLNPNINIVVVKVLLLRDRDSGPKVTG 283

Db 181 VVADESMWKFHGCADLEHYLLTLLATAARLYRHPSLNPNINIVVVKVLLLRDRDSGPKVTG 240
QY 284 NAALTLRNFCAWQKLNKVSOKHPEYWDTAIILFTRQDLCGATTCDTLGMADVGMCDPKR 343
Db 241 NAALTLRNFCAWQKLNKVSOKHPEYWDTAIILFTRQDLCGATTCDTLGMADVGMCDPKR 300
QY 344 SCSTVEDDGLPSAFTTAHELGHVFNPHDNVKNVCEVEFGKLRANHMSPPTLIODRANFW 403
Db 301 SCSTVEDDGLPSAFTTAHELGHVFNPHDNVKNVCEVEFGKLRANHMSPPTLIODRANFW 360
QY 404 SACSAAIITDFLDSDHGDCLLDQPSKPISLPEDLPAGSYTILSQOCELAFAVGSGSKPCPYMQ 463
Db 361 SACSAAIITDFLDSDHGDCLLDQPSKPISLPEDLPAGSYTILSQOCELAFAVGSGSKPCPYMQ 420
QY 464 YCTKLWCTGKAKQGMVCQTRHPPNADGTSCEGKLCILKAGCVERHNLNKH-----VDG 517
Db 421 YCTKLWCTGKAKQGMVCQTRHPPNADGTSCEGKLCILKAGCVERHNLNKHVVSSSQVIG 480
QY 518 SWAKWDPYGPCSRTCCGGVQLARQCTNPTPANGGKYCEGVVRKYRSCNLEPCPSSASGK 577
Db 481 SWAKWDPYGPCSRTCCGGVQLARQCTNPTPANGGKYCEGVVRKYRSCNLEPCPSSASGK 540
QY 578 SFREEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCLICRANGTGYYVLAKPVVD 637
Db 541 SFREEQCEAFNGYNHSTNRLTLAVAWPKYSGVSPRDKCLICRANGTGYYVLAKPVVD 600
QY 638 GTLCSPDSTSVCVQGCILKAGCNDNLGSKKRFDKCGVCGDNKCKVTGLFTKPMHGYN 697
Db 601 GTLCSPDSTSVCVQGCILKAGCNDNLGSKKRFDKCGVCGDNKCKVTGLFTKPMHGYN 660
QY 698 FVAIPAGASSIDIORGYKGLIGDDNYLALKNSQGYLLNGHFVWSAVERDLVWKGSL 757
Db 661 FVAIPAGASSIDIORGYKGLIGDDNYLALKNSQGYLLNGHFVWSAVERDLVWKGSL 720
QY 758 RYSGTGTAVESLQASRPILPLTVEVLSVCKMTPPRVRYSYFLPKPEPRDKSSHPRDPRG 817
Db 721 RYSGTGTAVESLQASRPILPLTVEVLSVCKMTPPRVRYSYFLPKPEPRDKSSHPRDPRG 780
QY 818 PSVLHNSVLSLSNOVEOPDDRRPPARWVAGSWGPCASCSGSLGKRAVDRCGSAGQRTVFA 877
Db 781 PSVLHNSVLSLSNOVEOPDDRRPPARWVAGSWGPCASCSGSLGKRAVDRCGSAGQRTVFA 840
QY 878 CDAHRPVTQACGEPCTWELSAWSPCKSCSGRGQRRLSKCV 921
Db 841 CDAHRPVTQACGEPCTWELSAWSPCKSCSGRGQRRLSKCV 884

RESULT 11

US-60-209-043-133

; Sequence 133, Application US/60209043

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL000641

; CURRENT APPLICATION NUMBER: US/60/209,043

; CURRENT FILING DATE: 2000-06-02

; NUMBER OF SEQ ID NOS: 232

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 133

; LENGTH: 884

; TYPE: PRT

; ORGANISM: HUMAN

US-60-209-043-133

Query Match 92.2%; Score 4758; DB 27; Length 884;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 878; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 44 GPDSGDOGLIFQITAFQEDFYHLHTPDQAFLAPAFSTHGLGVPLQGLTGSSDLRRCFY 103
|||||

Db 1 GPEDSGDQGLIFQITAFQEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRCFCY 60
QY 104 SCDVNAEPDPSAAVSLCGLGAGYGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVP 163
Db 61 SCDVNAEPDPSAAVSLCGLGAGYGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVP 120
QY 164 GPGSDPTSRGCVASGWNPAIIRALDPYKPRRAGFGESESRSSRRSRRAKRFVSPRYVETL 223
Db 121 GPGSDPTSRGCVASGWNPAIIRALDPYKPRRAGFGESESRSSRRSRRAKRFVSPRYVETL 180
QY 224 VYADESMVKFPGADLEHYLLATLTAARLYRHPISILNPINIVVVKVLLLRDSDGPKVTG 283
Db 181 VYADESMVKFPGADLEHYLLATLTAARLYRHPISILNPINIVVVKVLLLRDSDGPKVTG 240
QY 284 NAALTLRNFCAWOKLKNVSKDHPYWDTAILETRQDLGCAATTCDTLGMADVGMCDPKR 343
Db 241 NAALTLRNFCAWOKLKNVSKDHPYWDTAILETRQDLGCAATTCDTLGMADVGMCDPKR 300
QY 344 SCSTVEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEEFGKLRANHMSPTLIQIDRANPW 403
Db 301 SCSTVEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEEFGKLRANHMSPTLIQIDRANPW 360
QY 404 SACSAAITDFLDSHGDCLLDQPSKPIISLPEDLPGASYTILSQOCELAFVGSKPCPYMQ 463
Db 361 SACSAAITDFLDSHGDCLLDQPSKPIISLPEDLPGASYTILSQOCELAFVGSKPCPYMQ 420
QY 464 YCTKLWCTGKAKQWQVCOTRHPFWADGTSCEGKLCGLKACVERHNLNKHRR-----VDG 517
Db 421 YCTKLWCTGKAKQWQVCOTRHPFWADGTSCEGKLCGLKACVERHNLNKHRRVSSQVDG 480
QY 518 SWAKWDYPCSRCTCGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 577
Db 481 SWAKWDYPCSRCTCGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 540
QY 578 SPREQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 637
Db 541 SPREQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 600
QY 638 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDKSKCKVTGLFTKPMHGYN 697
Db 601 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDKSKCKVTGLFTKPMHGYN 660
QY 698 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLL 757
Db 661 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLL 720
QY 758 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPRVRYSYFLPKPEPKDPRG 817
Db 721 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPRVRYSYFLPKPEPKDPRG 780
QY 818 PSVLHNSVLSNVOEQDPPRPAWVAGSWGPCSCSGLOKRAVDCRSAGORTVPA 877
Db 781 PSVLHNSVLSNVOEQDPPRPAWVAGSWGPCSCSGLOKRAVDCRSAGORTVPA 840
QY 878 CDAHRPVETOACGPECTWELSAWSPCKSCGRGFORRLKCV 921
Db 841 CDAHRPVETOACGPECTWELSAWSPCKSCGRGFORRLKCV 884

RESULT 12

US-10-163-316-2
; Sequence 2, Application US/10163316
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; FILE REFERENCE: MP101-025P1RNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-316-2
Query Match 82.3%; Score 4248.5; DB 25; Length 823;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 792; Conservative 0; Mismatches 1; Indels 23; Gaps 1;
QY 1 MLLGLTLTAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPDSDGQGLIFQITAF 60
Db 1 MLLGLTLTAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPDSDGQGLIFQITAF 60
QY 61 QEDFYLHLTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRCFCYSGDVNAEPDPSFAAVSLC 120
Db 61 QEDFYLHLTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRCFCYSGDVNAEPDPSFAAVSLC 120
QY 121 GGLRGAFGYRGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVPGPGSDPTSRGCVASGW 180
Db 121 GGLRGAFGYRGAEEVVISPLNAPASAPAAQRNSQGAHLQRRGVPGPGSDPTSRGCVASGW 180
QY 181 NPAILRALDPYKPRRAGFGESESRSSRRSRRAKRFVSPRYVETILVVADESMVKFPGADLEH 240
Db 181 NPAILRALDPYKPRRAGFGESESRSSRRSRRAKRFVSPRYVETILVVADESMVKFPGADLEH 240
QY 241 YLLTLATAARLYRHPISILNPINIVVVKVLLLRDSDGPKVTGNAALTILRNFCAWOKKLN 300
Db 241 YLLTLATAARLYRHPISILNPINIVVVKVLLLRDSDGPKVTGNAALTILRNFCAWOKKLN 300
QY 301 KVS DKHPYWDTAILETRQDLGCAATTCDTLGMADVGMCDPKRSCSVIEDDGLPSAFTTA 360
Db 301 KVS DKHPYWDTAILETRQDLGCAATTCDTLGMADVGMCDPKRSCSVIEDDGLPSAFTTA 360
QY 361 HELGHVFNMPHDNVKVCVEEFGKLRANHMSPTLIQIDRANPWASACSAALITDFLDSHG 420
Db 361 HELGHVFNMPHDNVKVCVEEFGKLRANHMSPTLIQIDRANPWASACSAALITDFLDSHG 420
QY 421 DCLLDQPSKPIISLPEDLPGASYTILSQOCELAFVGSKPCPYMOYCTKLWCTGKAKQWVC 480
Db 421 DCLLDQPSKPIISLPEDLPGASYTILSQOCELAFVGSKPCPYMOYCTKLWCTGKAKQWVC 480
QY 481 QTRHFPWADGTSCEGKLCGLKACVERHNLNKHRR-----VDG 517
Db 481 QTRHFPWADGTSCEGKLCGLKACVERHNLNKHRRPTDIISPKQLLLLPNGLHTTQVDG 540
QY 518 SWAKWDYPCSRCTCGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 577
Db 541 SWAKWDYPCSRCTCGGVQLARRQCTNPTPANGGKYCEGVRYKRSCLNLEPCPSSASGK 600
QY 578 SPREQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 637
Db 601 SPREQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 660
QY 638 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDKSKCKVTGLFTKPMHGYN 697
Db 661 GTLCSPDSTSVCVQKCKIKAGCDNGLSKRFDKCGVCGGDKSKCKVTGLFTKPMHGYN 720
QY 698 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLL 757
Db 721 FYVAIPAGASSIDIRQRYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLL 780
QY 758 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPR 793
Db 781 RYSGTGTAVESLQASRPILPTVEVLSVGKMTTPR 816

RESULT 13

US-60-297-863-2
; Sequence 2, Application US/60297863
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: A Human Matrix Metalloproteinase and Uses

; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI2001-025P1(M)
; CURRENT APPLICATION NUMBER: US/60/297,863
; CURRENT FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-297-863-2

Query Match 82.3%; Score 4248.5; DB 27; Length 823;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 792; Conservative 0; Mismatches 1; Indels 23; Gaps 1;
QY 1 MLLGLITLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60
Db 1 MLLGLITLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFQITAF 60
QY 61 QEDFYHLTPDAQFLAPAFSTEHLGVLPQGLTGGSDLRRCFYSGDVNAEPDSFAAVSLC 120
Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVLPQGLTGGSDLRRCFYSGDVNAEPDSFAAVSLC 120
QY 121 GGLRGAFYRGAEVYISPLPNASAPAAQRNSOGAHLORRGVPGPGSDPTSRGCVASGW 180
Db 121 GGLRGAFYRGAEVYISPLPNASAPAAQRNSOGAHLORRGVPGPGSDPTSRGCVASGW 180
QY 181 NPAILRALDPYKPRRAGFSGRSRRSRGAKRFVSIPIRYVETLVVADESVMVRFHGADLEH 240
Db 181 NPAILRALDPYKPRRAGFSGRSRRSRGAKRFVSIPIRYVETLVVADESVMVRFHGADLEH 240
QY 241 YLTLTATAARLYRHSILNPINIVVVKVLLLRDROSGPKVTCGNAALTIRNFCAMOKKLN 300
Db 241 YLTLTATAARLYRHSILNPINIVVVKVLLLRDROSGPKVTCGNAALTIRNFCAMOKKLN 300
QY 301 KVSQKHEPYWDYTAILTRQDLGCATTCDTLGMADVGTCDPKRSCSVIEDDGLPSAFTTA 360
Db 301 KVSQKHEPYWDYTAILTRQDLGCATTCDTLGMADVGTCDPKRSCSVIEDDGLPSAFTTA 360
QY 361 HELGHVFNPHDNVNVKCEEVFGKLRANHMSPITLQIDRANPWSACSAAIITDFLDSGHG 420
Db 361 HELGHVFNPHDNVNVKCEEVFGKLRANHMSPITLQIDRANPWSACSAAIITDFLDSGHG 420
QY 421 DCLLDQSPKISLPEDLPGASYTLGACVGRHNLNKHRR-----VDG 517
Db 421 DCLLDQSPKISLPEDLPGASYTLGACVGRHNLNKHRR-----VDG 517
QY 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRRPTDIISPKQLLLRLPGLHTQVDG 540
Db 481 QTRHFPWADGTSCEGKCLKACVERHNLNKHRRPTDIISPKQLLLRLPGLHTQVDG 540
QY 518 SWAKWDPYPCSTCGGGVQLARRQCTNTPANGKYGKCVGRVYKRSCHLEPCPSASGK 577
Db 518 SWAKWDPYPCSTCGGGVQLARRQCTNTPANGKYGKCVGRVYKRSCHLEPCPSASGK 577
QY 541 SWAKWDPYPCSTCGGGVQLARRQCTNTPANGKYGKCVGRVYKRSCHLEPCPSASGK 600
Db 541 SWAKWDPYPCSTCGGGVQLARRQCTNTPANGKYGKCVGRVYKRSCHLEPCPSASGK 600
QY 578 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCLICRANGTGYFYVLAPKVV 637
Db 578 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCLICRANGTGYFYVLAPKVV 637
QY 601 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCLICRANGTGYFYVLAPKVV 660
Db 601 SFREOCEAFNGYNHSTNRLTLAVAWPKYSGVSPDKCLICRANGTGYFYVLAPKVV 660
QY 638 GTLCSPDSTSVCVQKRCIKAGCDNGLGSKKRDGKVCVCGDGNKSKKVTGFTKPMHGVN 697
Db 638 GTLCSPDSTSVCVQKRCIKAGCDNGLGSKKRDGKVCVCGDGNKSKKVTGFTKPMHGVN 697
QY 661 GTLCSPDSTSVCVQKRCIKAGCDNGLGSKKRDGKVCVCGDGNKSKKVTGFTKPMHGVN 720
Db 661 GTLCSPDSTSVCVQKRCIKAGCDNGLGSKKRDGKVCVCGDGNKSKKVTGFTKPMHGVN 720
QY 698 FVVAIPAGASSIDIRQYKGLIGDDNYLALNSCKYLLNGHFVVSAVERDLVVKGSLL 757
Db 698 FVVAIPAGASSIDIRQYKGLIGDDNYLALNSCKYLLNGHFVVSAVERDLVVKGSLL 757
QY 721 FVVAIPAGASSIDIRQYKGLIGDDNYLALNSCKYLLNGHFVVSAVERDLVVKGSLL 780
Db 721 FVVAIPAGASSIDIRQYKGLIGDDNYLALNSCKYLLNGHFVVSAVERDLVVKGSLL 780
QY 758 RYSGTGATVESIQASRPILPLETVEVLSVGMTPPR 793
Db 758 RYSGTGATVESIQASRPILPLETVEVLSVGMTPPR 793
QY 781 RYSGTGATVESIQASRPILPLETVEVLSVGMTPPR 816
Db 781 RYSGTGATVESIQASRPILPLETVEVLSVGMTPPR 816

RESULT 14
US-10-093-463-30
; Sequence 30, Application US/10093463
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zernusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 15:22:46 ; Search time 97 Seconds
(without alignments)
1674.819 Million cell updates/sec

Title: US-09-965-631-4

Perfect score: 5162

Sequence: 1 MLLGLTLFAGRTAGGSE.....DQNLHRKPQLDFCVLRPC 950

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 787003 seqs, 171007862 residues

Total number of hits satisfying chosen parameters: 787003

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	5162	100.0	950	6	US-10-391-364-77	Sequence 77, Appl
2	5162	100.0	950	7	US-60-453-135-14971	Sequence 14971, A
3	5162	100.0	950	7	US-60-453-050-14971	Sequence 14971, A
4	5117	99.1	952	6	US-10-311-035-11	Sequence 11, Appl
5	2486.5	48.2	967	6	US-10-369-779-18	Sequence 18, Appl
6	2486.5	48.2	967	7	US-60-452-680-23653	Sequence 23653, A
7	2486.5	48.2	967	7	US-60-453-135-14843	Sequence 14843, A
8	2486.5	48.2	967	7	US-60-453-050-14843	Sequence 14843, A
9	2486.5	48.2	967	7	US-60-455-444-8026	Sequence 8026, Ap
10	2486.5	48.2	999	6	US-10-369-779-31	Sequence 31, Appl
11	2485.5	48.1	951	6	US-10-369-779-29	Sequence 29, Appl
12	2485.5	48.1	951	6	US-10-381-793-3	Sequence 3, Appl
13	2485.5	48.1	968	6	US-10-369-779-19	Sequence 19, Appl
14	2485.5	48.1	968	6	US-10-391-364-82	Sequence 82, Appl
15	2482.5	48.1	967	7	US-60-440-068-134	Sequence 134, Appl
16	2474	47.9	967	6	US-10-369-779-20	Sequence 20, Appl
17	2457.5	47.6	951	6	US-10-369-779-28	Sequence 28, Appl
18	2138.5	41.4	890	5	US-09-949-002-394	Sequence 394, Appl
19	2138.5	41.4	890	6	US-10-369-779-27	Sequence 27, Appl
20	2135.5	41.4	821	5	US-09-724-676-80844	Sequence 80844, A
21	2135.5	41.4	821	5	US-09-724-676A-80844	Sequence 80844, A
22	2135.5	41.4	887	5	US-09-949-002-534	Sequence 534, Appl
23	2135.5	41.4	1004	7	US-60-455-444-8141	Sequence 8141, Ap
24	2130.5	41.3	890	5	US-09-724-676-67251	Sequence 67251, A
25	2130.5	41.3	890	5	US-09-724-676A-67251	Sequence 67251, A
26	1965	38.1	1916	6	US-10-274-639-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-391-364-77
; Sequence 77, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; FILE REFERENCE: MPI03-0190NMN
; CURRENT APPLICATION NUMBER: US/10/391,364
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-364-77

Query Match 100.0%; Score 5162; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLLGLILTAFAGRTAGGSEPREVVVPIRLDPDINGRRYYWRGPEDSGCUQLIFIQTAF	60
DB	1	MLLGLILTAFAGRTAGGSEPREVVVPIRLDPDINGRRYYWRGPEDSGCUQLIFIQTAF	60
QY	61	QEDFYLHLTPDAQFLAPASTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
DB	61	QEDFYLHLTPDAQFLAPASTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
QY	121	GGLRGAFYRGAEYVISPINASAPAAQNSOGAHLQRRYVGGPGSGDPTSRCGVASGW	180
DB	121	GGLRGAFYRGAEYVISPINASAPAAQNSOGAHLQRRYVGGPGSGDPTSRCGVASGW	180
QY	181	NPAILRALDPYKPRRAGFESRRRSRGRAKRFVSIPIRVETILVVADESWMKFHADLEH	240
DB	181	NPAILRALDPYKPRRAGFESRRRSRGRAKRFVSIPIRVETILVVADESWMKFHADLEH	240
QY	241	YLLTLTATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVGNAAALTLRNFCAWQKKLN	300
DB	241	YLLTLTATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVGNAAALTLRNFCAWQKKLN	300
QY	301	KVSDKHPEYWDTAILTFRODLGCATCTDGLMADVGTMDCKRSCSVIEDDGLPSAFTTA	360
DB	301	KVSDKHPEYWDTAILTFRODLGCATCTDGLMADVGTMDCKRSCSVIEDDGLPSAFTTA	360
QY	361	HELGHVFNMPHDNVKVEEVFGKLRANHMSPTLQIDRANPWSACSAAIITDFLDSGHG	420
DB	361	HELGHVFNMPHDNVKVEEVFGKLRANHMSPTLQIDRANPWSACSAAIITDFLDSGHG	420
QY	421	DCLLDPQSKPISLPEDLPGASYTLSQOCELAFCVGSKPCPYMOYCTKLWCTCKAKGQWVC	480
DB	421	DCLLDPQSKPISLPEDLPGASYTLSQOCELAFCVGSKPCPYMOYCTKLWCTCKAKGQWVC	480
QY	481	QTRHFPWADGTSCEGKCLKLGACVERHNLNKHRVDGWSWAKMDPYGPCSRCTCGGVQVLAR	540
DB	481	QTRHFPWADGTSCEGKCLKLGACVERHNLNKHRVDGWSWAKMDPYGPCSRCTCGGVQVLAR	540
QY	541	ROCTNPTPANGKYCEGVRVKYRSCNLEPCPSSASGKSPREOCFAFNYNHSTNRLTLA	600
DB	541	ROCTNPTPANGKYCEGVRVKYRSCNLEPCPSSASGKSPREOCFAFNYNHSTNRLTLA	600
QY	601	VAWPKYSGVSPRDKCLICRANGTYFYVLAPKVVYDGTLCSPDSTSVCVQKCIKAGCD	660
DB	601	VAWPKYSGVSPRDKCLICRANGTYFYVLAPKVVYDGTLCSPDSTSVCVQKCIKAGCD	660
QY	661	GNLGSKKRFDKCGVCGGDNKSKKVTGLTKPMHGNFYVVAIPAGASSIDIQRGYKGLI	720
DB	661	GNLGSKKRFDKCGVCGGDNKSKKVTGLTKPMHGNFYVVAIPAGASSIDIQRGYKGLI	720
QY	721	GDDNYLALKNISOGKYLINGHFVVSAYERDLVKGSLRLRYSGTAVESLQASRPFILEPLT	780
DB	721	GDDNYLALKNISOGKYLINGHFVVSAYERDLVKGSLRLRYSGTAVESLQASRPFILEPLT	780
QY	781	VEVLSVGKMTPPRVRSFYLPKPEPREDKSSHPKDPRGSPVLINSVLVLSNQVEQDDRRPP	840
DB	781	VEVLSVGKMTPPRVRSFYLPKPEPREDKSSHPKDPRGSPVLINSVLVLSNQVEQDDRRPP	840
QY	841	ARWVAGSWGPCSCSGGLQKRAVDCRGAGORTVPACDAAHRPVBTOACGPPCTWELLS	900
DB	841	ARWVAGSWGPCSCSGGLQKRAVDCRGAGORTVPACDAAHRPVBTOACGPPCTWELLS	900
QY	901	ANWSPKSCSGRGFORLSLKCVCVGHGBRLLARDQCNLHRKPOELDFCVLRPC	950
DB	901	ANWSPKSCSGRGFORLSLKCVCVGHGBRLLARDQCNLHRKPOELDFCVLRPC	950

RESULT 2

RESULT: 2
US-60-453-135-14971

```

US-60-433-133-149/1
; Sequence 14971, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: TAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISM

```


Db 841 ARWAGSWGPCSCGSGQLQRAYDCRGSAQRTVPACDAHRPVEQACGECPTWELS 900
Qy 901 AWSPCSKSCGRGFORRSKLCVGHGRLLRDOCNLHRKPQELDFCVLRPC 950
Db 901 AWSPCSKSCGRGFORRSKLCVGHGRLLRDOCNLHRKPQELDFCVLRPC 950
RESULT 3
US-60-453-050-14971
; Sequence 14971, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14971
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-14971
Query Match 100.0%; Score 5162; DB 7; Length 950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLLGILTLFAGTAGSGPEREVVPIRLDPDINGRRYYWKGPEDSGQGLFQITAF 60
Db 1 MLLGILTLFAGTAGSGPEREVVPIRLDPDINGRRYYWKGPEDSGQGLFQITAF 60
Qy 61 QEDFYHLTPDAQFLAPAFSTHGLVPLQGLTGSSDLRCFYSGDVNAEPDFAAVSLC 120
Db 61 QEDFYHLTPDAQFLAPAFSTHGLVPLQGLTGSSDLRCFYSGDVNAEPDFAAVSLC 120
Qy 121 GGLRGAFGYRGAEVVISPLPNASAPAAQNSQGAHLRLQRRVFGPGSGDPTSRGCVASGW 180
Db 121 GGLRGAFGYRGAEVVISPLPNASAPAAQNSQGAHLRLQRRVFGPGSGDPTSRGCVASGW 180
Qy 181 NPAILRALDPYKPRRAGFSGSRSSRRSGRAKRFVSIPIRYVETLVVADESXVKFHGADLEH 240
Db 181 NPAILRALDPYKPRRAGFSGSRSSRRSGRAKRFVSIPIRYVETLVVADESXVKFHGADLEH 240
Qy 241 YLLTLLATAARLYRHPISILAPINIVVVKVLLLRDROSGPKVTGNAALTLENFCAWOKKLN 300
Db 241 YLLTLLATAARLYRHPISILAPINIVVVKVLLLRDROSGPKVTGNAALTLENFCAWOKKLN 300
Qy 301 KVS DKHPEYWD TALFTRODLGCATTCDTLGMADVGTMCDPKRS CVIEDDGLPSAFTTA 360
Db 301 KVS DKHPEYWD TALFTRODLGCATTCDTLGMADVGTMCDPKRS CVIEDDGLPSAFTTA 360
Qy 361 HELGHVNMFDNKKVCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAITDFLDSHG 420
Db 361 HELGHVNMFDNKKVCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAITDFLDSHG 420
Qy 421 DCLLDQSKPISLPEDLPGASYTLISQCELAFGVSKPCPYMOYCKLWCTGKAKQMV 480
Db 421 DCLLDQSKPISLPEDLPGASYTLISQCELAFGVSKPCPYMOYCKLWCTGKAKQMV 480
Qy 481 QTRHFPWADTSGEGKCLKACVERHNLNKHRVDSNAKWBPYGPCSRTCGGVOLAR 540
Db 481 QTRHFPWADTSGEGKCLKACVERHNLNKHRVDSNAKWBPYGPCSRTCGGVOLAR 540
Qy 541 RQCTNPTPANGKYCEGVRYKRSCLNLEPCPSSASGKSFRECEAFNGYNHSTNLTILA 600
Db 541 RQCTNPTPANGKYCEGVRYKRSCLNLEPCPSSASGKSFRECEAFNGYNHSTNLTILA 600
Qy 601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVVDGTLCSPDSTSVCVQKCIKAGCD 660
Db 601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVVDGTLCSPDSTSVCVQKCIKAGCD 660

Db 601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVVDGTLCSPDSTSVCVQKCIKAGCD 660
Qy 661 GNIGSKRRDKCGVCGDNKSKKVTGLFTKPMHGVNFVVAIPAGASSIDIRORGYKGLI 720
Db 661 GNIGSKRRDKCGVCGDNKSKKVTGLFTKPMHGVNFVVAIPAGASSIDIRORGYKGLI 720
Qy 721 GDDNYLALKNSQGYKYLNGHFVVSAVERDLVVKGSLRLRYSGTGTAVESLQASRPILPLT 780
Db 721 GDDNYLALKNSQGYKYLNGHFVVSAVERDLVVKGSLRLRYSGTGTAVESLQASRPILPLT 780
Qy 781 VEVLVSGKMTPPRVRYSFYLPKREPREDKSSHPKDPGRPSVLHNSVLNSQVEQDDRRP 840
Db 781 VEVLVSGKMTPPRVRYSFYLPKREPREDKSSHPKDPGRPSVLHNSVLNSQVEQDDRRP 840
Qy 841 ARWAGSWGPCSCGSGQLQRAYDCRGSAQRTVPACDAHRPVEQACGECPTWELS 900
Db 841 ARWAGSWGPCSCGSGQLQRAYDCRGSAQRTVPACDAHRPVEQACGECPTWELS 900
Qy 901 AWSPCSKSCGRGFORRSKLCVGHGRLLRDOCNLHRKPQELDFCVLRPC 950
Db 901 AWSPCSKSCGRGFORRSKLCVGHGRLLRDOCNLHRKPQELDFCVLRPC 950
RESULT 4
US-10-311-035-11
; Sequence 11, Application US/10311035
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBBLEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; FILE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,9
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 952
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473089CD1
US-10-311-035-11
Query Match 99.1%; Score 5117; DB 6; Length 952;
Best Local Similarity 99.5%; Pred. No. 0;

Matches 947; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFOLITAF 60
DB 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFOLITAF 60

QY 61 QEDFYHLTLFPAQFLAPAFSTHGLVPLQGLTGGSDLRRCFYSGDVNAEPDSPAASVLC 120
DB 61 QEDFYHLTLFPAQFLAPAFSTHGLVPLQGLTGGSDLRRCFYSGDVNAEPDSPAASVLC 120

QY 121 GGLRGAFYRGAEYVISPPLNAPAAQNSOGAHLQRRGVPGSPDPTSRGCVASGW 180
DB 121 GGLRGAFYRGAEYVISPPLNAPAAQNSOGAHLQRRGVPGSPDPTSRGCVASGW 180

QY 181 NPAILRALDPYPRAGFESRRSRSSRAKRFVSIPIRYVETLVVADESXVKFPGADLEH 240
DB 181 NPAILRALDPYPRAGFESRRSRSSRAKRFVSIPIRYVETLVVADESXVKFPGADLEH 240

QY 241 YLLTLTAARLYRHPISILNPINIVVVKVLLLRDRSDGPKVTGNAAITLRFNFCAMQKLN 300
DB 241 YLLTLTAARLYRHPISILNPINIVVVKVLLLRDRSDGPKVTGNAAITLRFNFCAMQKLN 300

QY 301 KYSDKHPEYWDTAITLTFRODLGAGTCDPLGMADYGTMDCKPKRSCSVIEDDGLPSAFTTA 360
DB 301 KYSDKHPEYWDTAITLTFRODLGAGTCDPLGMADYGTMDCKPKRSCSVIEDDGLPSAFTTA 360

QY 361 HELGHVFNPHDNVNVKCEEVFGKLRANHMSPITLIQIDRANPWSACSAAITDFLDSHG 420
DB 361 HELGHVFNPHDNVNVKCEEVFGKLRANHMSPITLIQIDRANPWSACSAAITDFLDSHG 420

QY 421 DCLLDQSPKIPSLPDLPGAAYTLTQQCELAFGVSGKPCPYMOCYTKLWCTGKAKQWVC 480
DB 421 DCLLDQSPKIPSLPDLPGAAYTLTQQCELAFGVSGKPCPYMOCYTKLWCTGKAKQWVC 480

QY 481 QTRHFPWADGTSCGEGKCLKACVERHNLNKHRYDGVGSAKWDPYPCSKTCGGVQQLAR 540
DB 481 QTRHFPWADGTSCGEGKCLKACVERHNLNKHRYDGVGSAKWDPYPCSKTCGGVQQLAR 540

QY 541 RQCTNTPANGKCYCBGVKRYRSCNLEPCSSASGKSFREQCCEAFNGYNHSTNRLTLA 600
DB 541 RQCTNTPANGKCYCBGVKRYRSCNLEPCSSASGKSFREQCCEAFNGYNHSTNRLTLA 600

QY 601 VAWPKYSGVSPDRCKLICRANGTGYFYVLAPK-VVDGTLCSPDSTSVCVQKCIKAGC 659
DB 601 VAWPKYSGVSPDRCKLICRANGTGYFYVLAPK-VVDGTLCSPDSTSVCVQKCIKAGC 660

QY 660 DGNLGSKKRFDKCGVCGGDNKCKKVTGLFTKPMHGYNEVVAIPAGASSIDIRQRYKGL 719
DB 660 DGNLGSKKRFDKCGVCGGDNKCKKVTGLFTKPMHGYNEVVAIPAGASSIDIRQRYKGL 720

QY 720 IGDNDYLLKNSQGYLLNGHFVVSVERDLVVKGSLRLRYSGTGAVESLQASRPILP 779
DB 720 IGDNDYLLKNSQGYLLNGHFVVSVERDLVVKGSLRLRYSGTGAVESLQASRPILP 780

QY 780 TVEVLSVGMKTPRVRYSYFLPEKEDKSSHPKDP-RGPSVLHNSVLSLNSQVQPDPR 838
DB 780 TVEVLSVGMKTPRVRYSYFLPEKEDKSSHPKDP-RGPSVLHNSVLSLNSQVQPDPR 840

QY 839 PPARWAGSWGPCASGSGLOKRAVDRCGSAQRTVPACDAARHPVETOACGECPTWE 898
DB 839 PPARWAGSWGPCASGSGLOKRAVDRCGSAQRTVPACDAARHPVETOACGECPTWE 900

QY 899 LSAMSPCSKSGRGFORRLKCVGHGRLRLARQCNLHRKPOELDFCVLRPC 950
DB 901 LSAMSPCSKSGRGFORRLKCVGHGRLRLARQCNLHRKPOELDFCVLRPC 952

RESULT 5

US-10-369-779-18
; Sequence 18, Application US/10369779
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Sharma, Rahul

; APPLICANT: Ramanathan, Chandra
; APPLICANT: Westphal, Ryan
; APPLICANT: Feder, John
; TITLE OF INVENTION: NOVEL HUMAN ADAM-TS POLYPEPTIDE CONTAINING A DISINTEGRIN AND
; TITLE OF INVENTION: METALLOPROTEASE DOMAIN WITH THROMBOSPONDIN TYPE-1 MOTIFS,
; TITLE OF INVENTION: ADAM-TS-SI
; FILE REFERENCE: D0204 NP
; CURRENT APPLICATION NUMBER: US/10/369,779
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/358,151
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-779-18

Query Match 48.2%; Score 2486.5; DB 6; Length 967;
Best Local Similarity 49.0%; Pred. No. 2.3e-174;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDSDGGLIFOLITAF 60
DB 36 LLLAAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTRLRLHAF 81

QY 61 QEDFYHLTLFPAQFLAPAFSTHGLVPLQGLTGGSDLRRCFYSGDVNAEPDPSF 114
DB 82 DQQLDLRLPDSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136

QY 115 AAVSLCGLRGAFYGAAYVISPLPNAS---APAAQNSQGA-----HLLQ---RRGVPGG 165
DB 137 AALSCEGVRGAYLLGEAYFIQPLPAASERLATAPGEEKPPAPLQFHLRLRRNRQDVG 196

QY 166 PSG-----DPTSRC-----GVASG--WNPAITLRALDPYKPRRAGEGSRSSRRS 207
DB 197 TCGVVVDPEPRTGAETDEDETEGEDEGPQWS-----PODPALQGVGP-IIGS 247

QY 208 GRAKRFYSIPRYVETLVVADESXVKFPGADLEHYLLTLTAARLYRHPISILNPINIVV 267
DB 248 IRKKRFVSSHRYVETLVVAQNSAEFGSLKHYLLTFLSVAAARLYKHPISIRNSVSLVV 307

QY 268 KVLRLDRDQSPKVTGNAAITLRFNFCAMQKLNKVSQKHPEYWDTAITLTFRODLGATTC 327
DB 308 KILVIHQEGKPEVTSNAALTLRNFCNMWQKHPPSDAEHYDTAIFLTRQDLCSQTC 367

QY 328 DTLGMADVTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKCEEVFGKLRAN 387
DB 368 DTLGMADVTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKCEEVFGKLRAN 427

QY 388 HMSGPTLIQIDRANPWSACSAAITDFLDSHGDCLLDQPSKIPSLPDLPGAAYTLTQQ 447
DB 428 HMASMLSLNDHSQPSWSPSAYSMITSLDNGHGBCLMDKPNQPIQLPGLDPTSYDANRQ 487

QY 448 CELAFGVSGKPCPYMOCYTKLWCTGKAKQWVCOTRHPWADGTSCGEGKCLKAGACVE 506
DB 488 CQFTFGEDSKHCPDAASTCTLWCTGTSGVVLVCQTKHFPWADGTSCGEGKWCINGKCVN 547

QY 507 RHNLNKH---RVDGSAKWADPYPCSKTCGGVQQLARQCTNPTPANGKCYCBGVKRYK 563
DB 548 KTD-RKHEDTFPHSGWGMWGPWGDCTSCGGVQVYTRCEDNPPVKNKGKCEGRVRYR 606

QY 564 SCNLEPCPSSASGKSFREQCCEAFNGYNHSTNRLTLAVAWPKYSGVSPDRCKLICRAN 623
DB 607 SCNLEPCPDN-NGKTFREQCCEAFNEFSKASFGSGPAVEWIPKYAGVSPDRCKLICQAK 665

QY 624 GTGYFYVLAPKVVDTLCSPDSTSVCVQKCIKAGCQCNLGRKPKDCKGVCQGDNKSCK 683
DB 666 GIGYFFVLQPKVVDGTPCSPDSTSVCVQGVQKAGCDRIIDSKKKFKDKCGGNGSGICK 725

QY 684 KVTGLTKPMHGYNEVVAIPAGASSIDIRQRYKGLIGDNDYLLKNSQGYLLNGHFV 743
DB 935 KVTGLTKPMHGYNEVVAIPAGASSIDIRQRYKGLIGDNDYLLKNSQGYLLNGHFV 743

```
Db 726 KTSQSVTSKPGYHDIITPTGATNIEVKQNRGRNNGSFLAIIKAADGYIILNGDYTL 785
QY 744 SAVERDLVVKGLRYSGTGAVESLOASRPILPTVEVLSVGKMTPPKRVYRFFYLPKE 803
Db 786 STLEODIMYKGVLYRSGSSAALERISFSLKEPLTIOVLTVGNALRPKIKYIFYVKK 845
QY 804 PREDKSSHPKDPGRPSVILHNSVLNSNOVEQDPRPARWAGSWGPCASCGSLQKRA 863
Db 846 ----KES-----FNAIPTS-----AWIEEWGECSSCELGWQRL 878
QY 864 VDCRSAGQRTVPACDAH--RPVETOACGE-PCPTWELSAWSPCSKSGHGFORSILK 920
Db 879 VECRDINGO---PASECAKEVKPASTRPCADHPCPQWLGEMSSCKTCGKGYKRSILK 935
QY 921 VGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKPKHFIDECTMAEC 966

RESULT 6
US-60-452-680-23653
; Sequence 23653, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23653
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-23653

Query Match 48.2%; Score 2486.5; DB 7; Length 967;
Best Local Similarity 48.5%; Pred. No. 2.3e-174;
Matches 483; Conservative 153; Mismatches 249; Indels 111; Gaps 22;

QY 1 MLLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPDSDGGLIFQITAF 60
Db 36 LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTLRLLHAF 81
QY 61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSDLRRCFYSGDVNAEPDSF 114
Db 82 DQQLDLRLPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAGYRGAEYVISPPLNAS---APAAQRNSOGA---HLLQRRGVPGGPS 167
Db 137 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAPGKPPAPLQFHLRR-----NRQ 191
QY 168 GDPTSRGCV-----ASGNPAILRALDPYKPRRAGFGESR 202
Db 192 GDVGTCGVVDDEPRPTKAETDEDECTEGEDEGAQNS-----PQDFALQGVGP- 242
QY 203 SRRRSGRAKRFVSPRYVETLVVADESVMVKFGADLEHYLTTLATAARLYRHPSILNPI 262
Db 243 TGTGIRKKRFVSSHYVETMLVADQSMAEFHGSGIKHYLLTLFVSAAARLYKHPISRSV 302
QY 263 NTVVVKVLLLRDRDSKPVGTGNAALTNRFCAWOKKLNKVSQKHPYWDYTAILETRDLC 322
Db 303 SLVVVKILVIHQEGPEVTSNAULTLRNFCNQKHPPSDROAEHYDIAILLTRDLC 362
QY 323 GATTCDTLGMADVTCMPKRSVCVIEDDGLPSAFETTAHELGHVFNPMHNVKVEEYFG 382
Db 363 GSQTCDTLGMADVTCVCDPSRSCVIEDDGLQAETTAHELGHVFNPMHDDAKQCASLNG 422
QY 383 KLRANHMSPTLIQIDRANPWSACAAIITDFLDGSHGDCLLDQSPKISIPEDLPGASY 442
Db 423 VNQDSHMLATSLNLDHSPQWSPSCSAYMTTSFLDNGHGECLMDKPNQPIQLPGDLPGTSY 482
```

```
QY 443 TLSQOCELAFGVGSKPCP-YMOYCTKLWCTKAKGQMVQCTRHFPWADGTSCGEGKLCLK 501
Db 483 DANRQCOFTGEDSKDPCDAASTCTLWCTGTSGVLVCQTKHFPWADGTSCGEGKWCIN 542
QY 502 GACVERHNLNKH---RVDGSAKWADYGPSCSRCTCGGQVLARRQCTNPPTPANGKCYCEG 558
Db 543 GKCVNKTID-RKHFDTPFHGSGWMPGWDGCSRTCGGVQYTMRECDNPYPKNGSKYCEG 601
QY 559 RVKVRSCNLEPCSSAGSGFREEOCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCL 618
Db 602 RVIRSCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKTAGVSPKDRCKL 660
QY 619 ICRANETGYFVYLPAPKVVVDGTLCSPDSTSVCOGKCIKAGCCDNLGSKRFDKCGVCGD 678
Db 661 ICQAKGIGYFVLPQKVVVDGTPCSPDSTSVCOGQCVKAGCDRIIDSKKKFDKCGVCGN 720
QY 679 NKCKKVTGLFTKPMHGYNFVVAIPAGASSIDTRQRYGKGLIGDDNYLALKNSQGYLLN 738
Db 721 GSTCKKISGVTSAKPGYHDIITPTGATNIEVKQNRGRNNGSFLAIIKAADGYIILN 780
QY 739 GHFVSAVERDLVVKGLRYSGTGAVESLOASRPILPTVEVLSVGKMTPPRVYRSP 798
Db 781 GDTLSTLEODIMYKGVLYRSGSSAALERISFSLKEPLTIOVLTVGNALRPKIKYIY 840
QY 799 YLPKEPREDKSSHPPKDPGRPSVILHNSVLNSNOVEQDPRPARWAGSWGPCASCGSG 858
Db 841 FVKKK-----KES-----FNAIPTS-----AWIEEWGECSSCELG 873
QY 859 LQKRAVDCRSAGQRTVPACDAH--RPVETOACGE-PCPTWELSAWSPCSKSGRGFOR 915
Db 874 WQRLVECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEMSSCKTCGKGYK 930
QY 916 RSLKCVCHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 931 RSLKCLSHDGGVLSHESCDPLKPKHFIDECTMAEC 966

RESULT 7
US-60-453-135-14843
; Sequence 14843, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14843
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-14843

Query Match 48.2%; Score 2486.5; DB 7; Length 967;
Best Local Similarity 48.5%; Pred. No. 2.3e-174;
Matches 483; Conservative 153; Mismatches 249; Indels 111; Gaps 22;

QY 1 MLLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYWRGPDSDGGLIFQITAF 60
Db 36 LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTLRLLHAF 81
QY 61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSDLRRCFYSGDVNAEPDSF 114
Db 82 DQQLDLRLPDSFSLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAGYRGAEYVISPPLNAS---APAAQRNSOGA---HLLQRRGVPGGPS 167
Db 137 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAPGKPPAPLQFHLRR-----NRQ 191
```

```
QY 168 GDPTSRGCV-----ASGNPAILRALDPYKPRRAGFESR 202
Db 192 GVGGTGCVVDDPRPTGKAETEDEGTEDEGAQWS-----PODPALQGVGP- 242
QY 203 SRRSGRAKRFVSIPIRYVETLVVADESVMKFGADLEHLLTLLATAARLYRHPSTILNPI 262
Db 243 TGTGSTRKRRFVSSHRYVETMLVADOSMAEFHSGGLKHYLLTLFVSAARLYKHPSTIRNSV 302
QY 263 NIIVVKVLLLRDRSDSGPKVTGNAAALTLRNFCAWKLNKVSQKHPEYWDTAIFLTRQDLC 322
Db 303 SLVVVKILVHDQKPEVTSNAAALTLRNFCAWKLNKVSQKHPEYWDTAIFLTRQDLC 362
QY 323 GATCTDLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHNVKVEEVFG 382
Db 363 GSQCTDLGMADVTCVCDPSRSCSVIEDDGLQAFAFTTAHELGHVFNMPHNDKACASLNG 422
QY 383 KLRANHMWSPTLQIDRANPWSACSAIITDFLDSHGDCCLDOPSKPISLPEDLPASV 442
Db 423 VNQDSHMASMLNLDHSQWSPSCSAIYITSLDNGHGECLMDKQNPITLPGDLPGETSY 482
QY 443 TLSQOCELAFVGSKPCP-YMQYCTKLWCTGKAKGOMVCQTRHFPWADGTSCEGKLCILK 501
Db 483 DANROQOFTGEDSKHCPDAASTCTLWCTGSGVLVQCTKHFPAWADGTSCEGKWCIN 542
QY 502 GACVERHNLKH---RVDGSNAKWDYPGCSRTCCGGVQLARROCTNPTPANGKVCYEGV 558
Db 543 GKCVNKT-D-RKHFTDPFHSGWMMGPMWDCSRTCCGGVQYTMRECDNPVFNKGKCYCEG 601
QY 559 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSVSPRDKCKL 618
Db 602 RVRYRSCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSPAVEWIPKYAGVSPKDRCKL 660
QY 619 ICRANGTGIFYVLAPKVVDGTLCSPDSTSVCGQKCIKAGDGNLGSKKRFDKCGVCGGD 678
Db 661 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFKDKCGVCGGN 720
QY 679 NKSCCKVTGLTKPMHGYNFVAIPAGASSIDIRQYKGLIGDDNYLALKNQSGKYLIN 738
Db 721 GSTCKKISGVSATKPGYHDIITPTGATNIEVKQRNQRNNGSFLAIKAADGTIYL 780
QY 739 GHFVVSVERDLVVKGLSLLRYSGTGAVESLQASRPILPLETVESLVGKMTPPRVYRSF 798
Db 781 GDYTLSTLEQDIYKGVWLYRSGSSAALERIRSFSPKLEPLTIQVLTGVNALRPKIKTY 840
QY 799 YLPKEPREDKSSHPKDPGSPVLHNSVLSNQVEQDPRPARVAGSWGPCSCSGG 858
Db 841 FVKKK----KES-----FNAIPTFS-----AWVIEEWGECRSKCELG 873
QY 859 LQKRAVDCRSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSCGRGFQR 915
Db 874 WQRLVECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCSTCGKGYKK 930
QY 916 RSLKCVGHGGRLLARDQCNLHRRKPOE-LDFCVLRPC 950
Db 931 RSLKCLSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966
```

RESULT 8

```
US-60-453-050-14843
; Sequence 14843, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453, 050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14843
; LENGTH: 967
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-60-453-050-14843
Query Match 48.2% Score 2486.5 DB 7 Length 967;
Best Local Similarity 48.5% Pred No. 2.3e-174;
Matches 483; Conservative 153; Mismatches 249; Indels 111; Gaps 22;
QY 1 MLLGIITLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPDGSGDGLFIQITAF 60
Db 36 LLLAAALAVSDALGRPSEDEELVVP-ELE-----RAP-----GHGTTLRRLHAF 81
QY 61 QEDFYHLHTPAQFLAPAFSTEHLG-----VPLQGLTGGSSDLRRCFYSGDVNAEPDSF 114
Db 82 DQQLDLELRPDSSFLAPGFTLQNYGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136
QY 115 AAVSLGGLRGAFGARGAEYVISPLPNAS---APAAQRNSOGA---HLQRRGVPGGPS 167
Db 137 AALSCEGVGAFYLLGEAFIQPLPAASERLATAAAGEKPPAPLOPHLLRR-----NRQ 191
QY 168 GDPTSRGCV-----ASGNPAILRALDPYKPRRAGFESR 202
Db 192 GVGGTGCVVDDPRPTGKAETEDEGTEDEGAQWS-----PODPALQGVGP- 242
QY 203 SRRSGRAKRFVSIPIRYVETLVVADESVMKFGADLEHLLTLLATAARLYRHPSTILNPI 262
Db 243 TGTGSTRKRRFVSSHRYVETMLVADQSMABFHSGGLKHYLLTLFVSAARLYKHPSTIRNSV 302
QY 263 NIIVVKVLLLRDRSDSGPKVTGNAAALTLRNFCAWKLNKVSQKHPEYWDTAIFLTRQDLC 322
Db 303 SLVVVKILVHDQKPEVTSNAAALTLRNFCAWKLNKVSQKHPEYWDTAIFLTRQDLC 362
QY 323 GATCTDLGMADVTCMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHNVKVEEVFG 382
Db 363 GSQCTDLGMADVTCVCDPSRSCSVIEDDGLQAFAFTTAHELGHVFNMPHNDKACASLNG 422
QY 383 KLRANHMWSPTLQIDRANPWSACSAIITDFLDSHGDCCLDOPSKPISLPEDLPASV 442
Db 423 VNQDSHMASMLNLDHSQWSPSCSAIYITSLDNGHGECLMDKQNPITLPGDLPGETSY 482
QY 443 TLSQOCELAFVGSKPCP-YMQYCTKLWCTGKAKGOMVCQTRHFPWADGTSCEGKLCILK 501
Db 483 DANROQOFTGEDSKHCPDAASTCTLWCTGSGVLVQCTKHFPAWADGTSCEGKWCIN 542
QY 502 GACVERHNLKH---RVDGSNAKWDYPGCSRTCCGGVQLARROCTNPTPANGKVCYEGV 558
Db 543 GKCVNKT-D-RKHFTDPFHSGWMMGPMWDCSRTCCGGVQYTMRECDNPVFNKGKCYCEG 601
QY 559 RVKYRSCNLEPCPSSASGKSFREOCEAFNGYNHSTNRLTLAVAWVPKYSVSPRDKCKL 618
Db 602 RVRYRSCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSPAVEWIPKYAGVSPKDRCKL 660
QY 619 ICRANGTGIFYVLAPKVVDGTLCSPDSTSVCGQKCIKAGDGNLGSKKRFDKCGVCGGD 678
Db 661 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFKDKCGVCGGN 720
QY 679 NKSCCKVTGLTKPMHGYNFVAIPAGASSIDIRQYKGLIGDDNYLALKNQSGKYLIN 738
Db 721 GSTCKKISGVSATKPGYHDIITPTGATNIEVKQRNQRNNGSFLAIKAADGTIYL 780
QY 739 GHFVVSVERDLVVKGLSLLRYSGTGAVESLQASRPILPLETVESLVGKMTPPRVYRSF 798
Db 781 GDYTLSTLEQDIYKGVWLYRSGSSAALERIRSFSPKLEPLTIQVLTGVNALRPKIKTY 840
QY 799 YLPKEPREDKSSHPKDPGSPVLHNSVLSNQVEQDPRPARVAGSWGPCSCSGG 858
Db 841 FVKKK----KES-----FNAIPTFS-----AWVIEEWGECRSKCELG 873
QY 859 LQKRAVDCRSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSCGRGFQR 915
Db 874 WQRLVECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCSTCGKGYKK 930
QY 916 RSLKCVGHGGRLLARDQCNLHRRKPOE-LDFCVLRPC 950
Db 931 RSLKCLSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966
```

Db 931 RSLKLSHDGVLSHESCDPLKKPKHFIDFCTMAEC 966

RESULT 9
US-60-455-444-8026
; Sequence 8026, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01455
; CURRENT APPLICATION NUMBER: US/60/455.444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8026
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-8026

Query Match 48.2%; Score 2486.5; DB 7; Length 967;
Best Local Similarity 48.5%; Pred. No. 2.3e-174;
Matches 483; Conservative 153; Mismatches 249; Indels 111; Gaps 22;

QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDESGOGLIFQITAF 60
Db 36 LLLLAALLAVSDALGRSEDEELVVP-ELE-----RAP---GHUTTRLRHAF 81
QY 61 QEDFYHLTPDAOFLAPAFSTEHLG-----VPLQGLTGSSDLRRCFYSGDVNAEPDSF 114
Db 82 DQDLLELRDSSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFYRGAEYVISPPLPNAS---APAAQNSQGA-----HLLQ---RMGVPGG 165
Db 137 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLQFHLRR-----RAP 191
QY 168 GDTSRGCV-----ASGNPALLRALDPYKPRAGFGESE 202
Db 192 GDVGTGCVVDDEPRPTGKAETDEDEGEDEGAQWS-----ETDLAHCFYSGTVNGDPSSA 242
QY 203 SRRRSGRARFYISPRYVETLVVADESMVYKFGADLEHLLTLLATAARLYRHPSILNPI 262
Db 243 TGTGSIRKKRFVSSHRYVETLVVADESMVYKFGADLEHLLTLLATAARLYRHPSILNPI 302
QY 263 NIIVYVLLRRDRSGPKVTGNAALTNRFCAMQKLNKVDKHPYWDITAILFTRQDL 322
Db 303 SLVVYKILVIHDEQKPEVTSNAALTNRFCAMQKLNKVDKHPYWDITAILFTRQDL 362
QY 323 GATTCDTLGMADVTMCDPKRSCSVIEDDGLPSAETTAHELGHVFNMPHDNVKYCEEVFG 382
Db 363 GSQCDTLGMADVTMCDPKRSCSVIEDDGLPSAETTAHELGHVFNMPHDNVKYCEEVFG 422
QY 383 KLRANHMSPPTLIQIDRANPWSACAAITDPLDSGHGDCLLDQPKSPISPELDPGASY 442
Db 423 VNQDSHMMASMLNLDHSPWSPCSAYMTSFLDNGHCEGLMDKPNQIQLPGLDPTSY 482
QY 443 TLSQCELAGFVGSFPCP-YMQYCHKLWCTGKAKQCMVQCTRHFPWADTSCGEGKCLCK 501
Db 483 DANROQFTFGEDSKHCPDAASTCTLMCTGTSGVLVQCOTKHPFWADTSCGEGKWCIN 542
QY 502 GACVERHNLNKE---RVDSMAKWPDPYPCSRCTCGGVQLARQCTNPTPANGKYCEGV 558
Db 543 GKCVNKT-DKHFDPFHSGWGMWGPWGDCSRTCGGVQVYTMRECDNPVKNNGKYCEGK 601
QY 559 RYKYSRNLPECPSSAKSFREOCEAFNGYNHSTNRLTLAVANVPKYSVSPDKCKL 618
Db 602 RYKYSRNLPECPDN-NKTFREOCEAHNEESKAFSGGPAVEWIPKYGSPDKCKL 660
QY 619 ICRANGTYFYVLAQVVDGTLCSPDSTSVQVQKICIKAGCDGNLGSKKRFDCGVCGD 678
Db 661 ICQAKGIGYFVLAQVVDGTLCSPDSTSVQVQKICIKAGCDGNLGSKKRFDCGVCGD 720

QY 679 NKCKKVTGLTTPMHGYNFVVAIPAGASSIDIRQYKGLIGDDNYLALKNSCGKYLLN 738
Db 721 GSTCKKISGVSYSAPGYHDIITPTGATNIEVKQRNQRNNGSFLAKRAADGTYILN 780
QY 739 GHFVYSAVERDLVVKGLRLRYSGTGAVESLQASRPILPLTVEVLVSGKMTFPRVYSF 798
Db 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALIRSFSPLEPLTIQVLTGVNALREKIKTY 840
QY 799 YLPKEPREDKSSHPKDPGRPSVLNSVLSNQVEQDDPRPPAKWVAGSGPCASCGSG 858
Db 841 FVKKK-----KES-----FNAIPTE-----AWVTEEWGECSSKSELG 873
QY 859 LOKRAVDGRSAGORTVPACDAH--RPVETQACGE-PCPTWELSAWSPCKSCGRGFR 915
Db 874 WQRLVECRDINGO---PASECAKEVPASTRCPADHPCPOMQLGEWSSCSKTCGKGYYK 930
QY 916 RSLKCVGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 931 RSLKLSHDGVLSHESCDPLKKPKHFIDFCTMAEC 966

RESULT 10
US-10-369-779-31
; Sequence 31, Application US/10369779
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Sharma, Rahul
; APPLICANT: Ramathanan, Chandra
; APPLICANT: Westphal, Ryan
; APPLICANT: Feder, John
; TITLE OF INVENTION: NOVEL HUMAN ADAM-TS POLYPEPTIDE CONTAINING A DISINTEGRIN AND
; TITLE OF INVENTION: METALLOPROTEASE DOMAIN WITH THROMBOSPONDIN TYPE-1 MOTIFS,
; FILE REFERENCE: ADAM-TS-SI
; CURRENT APPLICATION NUMBER: US/10/369,779
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/358,151
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 31
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-779-31

Query Match 48.2%; Score 2486.5; DB 6; Length 999;
Best Local Similarity 49.0%; Pred. No. 2.4e-174;
Matches 486; Conservative 154; Mismatches 250; Indels 101; Gaps 24;

QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDESGOGLIFQITAF 60
Db 68 LLLLAALLAVSDALGRSEDEELVVP-ELE-----RAP---GHUTTRLRHAF 113
QY 61 QEDFYHLTPDAOFLAPAFSTEHLG-----VPLQGLTGSSDLRRCFYSGDVNAEPDSF 114
Db 114 DQDLLELRDSSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 168
QY 115 AAVSLCGGLRGAFYRGAEYVISPPLPNAS---APAAQNSQGA-----HLLQ---RMGVPGG 165
Db 169 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLQFHLRRNRQGVGG 228
QY 166 PSG-----DPTSR-----GVASG--WNPAIRALDPYKPRRAGFGSGSKSRRS 207
Db 229 TCGVVDDEPRPTGKAETDEDEGEDEGAQWS-----ETDLAHCFYSGTVNGDPSSA 279
QY 208 GRARFYSIPRYVETLVVADESMVYKFGADLEHLLTLLATAARLYRHPSILNPI 267
Db 280 IRKRFRVSSHRYVETLVVADESMVYKFGADLEHLLTLLATAARLYRHPSILNPI 339
QY 268 KVLRLRRDRSGPKVTGNAALTNRFCAMQKLNKVDKHPYWDITAILFTRQDLGATTC 327

Db 340 KILVIHDEQKPEVTSNAALTLRNFNCWQKQHNPPSDRDAEHYDTAILFTIRQDLQSGQTC 399
QY 328 DTLMADVGTMCDDPKKSCSVIEDGLPSAFTTAHELGHVFNMPHDNVKVCVEVFGKLRAN 387
Db 400 DTLMADVGTMCDDPKKSCSVIEDGLPSAFTTAHELGHVFNMPHDNVKVCVEVFGKLRAN 459
QY 388 HMMSPFLIODIRANPWSACSAATITDPLDSHGDCCLLDQPSKPISTLPEDLPASYSYLSQO 447
Db 460 HMMASMLNSHDSQSPSCSAIMYTSFLDNGHGECCLMDKPNQIPQDPLPGTSDANRQ 519
QY 448 CELAFGVGSKPCP-YMQYCTKLMCTKAKQWQVCTRHFPWADGTSCGSGKCLKGACV 506
Db 520 CQFTFGEDSKHCPDAASTCTLWCTSGVLVCTKHFPWALGTSGEGKWCINGKCVN 579
QY 507 RHLNKH---RVDGSNAKWDYPGCSRTCGGGVQLARRQCTNTPANGKAYCEGVRYK 563
Db 580 KTD-RKHFTPTFHSGMGMWGPWDCSRTCGGVQVYTMRECDNPVKNKGKCYCBGKRVYR 638
QY 564 SCNLEPCPSASCSFREOCEAFNGYNHSTNRLTLAVAVPKYSGVSPKDKCLICRAN 623
Db 639 SCNLEDCPDN-NGKTFREOCEAHNEFSKASFGSGGPAVEWIPKAYAGVSPKDRCKLICQAK 697
QY 624 GTGYFVLAPKVVDGTLCSPDSTSVCGQKCIKAGCDGNLGSKKRFDKCGVCGGDNKSK 683
Db 698 GIGYFVLQPKVVDGTPCSPDSTSVCGQCYKAGCDRIIDSKKFDKCGVCGGNGSTCK 757
QY 684 KVTGLTKPMHGNFVVAIPAGASSIDIRQGYKGLIGDDNYALNKSQKYLINGHFVV 743
Db 758 KISGVSATKAPGVDHITITPTGATNIEVKORNGSRNNGSFLAIKAADSYTLNGDYTL 817
QY 744 SAVERDLVKSLLRYSGTGTAVESLOASRPILPLETLVEVLSVCKMTPPVRYVSFYLPR 803
Db 818 STLEQDMTKGVVLYRSGSAALEIRSFSPKPELTITQVLTGNALRPKRYTFVKK 877
QY 804 PREDKSHHPDRGPRSVLNSVLSLNOVEQPDPRPPARWAGSWGSPCSASCSGLQKRA 863
Db 878 ---KES-----FNAIPTFS-----ANVIEEWGECSSKSELGHQRL 910
QY 864 VDCRGSAQORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSGRGQRSLKC 920
Db 911 VECRDINGQ---PASECAKEVKPASTRPCADHPQWOLGESSCSKTCGKYKRSLKC 967
QY 921 VGHGGILLARDONLHRKPOE-LDFCVLRPC 950
Db 968 LSHDGGVLSHESCDPLKKPKHFDCTMAEC 998

RESULT 11

US-10-369-779-29
; Sequence 29, Application US/10369779
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Sharma, Rahul
; APPLICANT: Ramanathan, Chandra
; APPLICANT: Westphal, Ryan
; APPLICANT: Feder, John
; TITLE OF INVENTION: NOVEL HUMAN ADAM-TS POLYPEPTIDE CONTAINING A DISINTEGRIN AND
; TITLE OF INVENTION: METALLOPROTEASE DOMAIN WITH THROMBOSPONDIN TYPE-1 MOTIFS,
; FILE OF INVENTION: ADAM-TS-SI
; FILE REFERENCE: D0204 NP
; CURRENT APPLICATION NUMBER: US/10/369,779
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/358,151
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: patentin version 3.2
; SEQ ID NO 29
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-369-779-29

Query Match

48.1%; Score 2485.5; DB 6; Length 951;

RESULT 12

Best Local Similarity 48.0%; Pred. No. 2.7e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;
QY 1 MLLGLITLAFAGTAGG--SEPEREVVVPPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI 57
Db 20 LLLLASITMLLCARGAGHPTDEELVLP-SLE-----RAPHDSTTIRL--FL 66
QY 58 TAPQEDFYHLPTDPAQFLAPASTHEHLGVLPLQGLTGS-----SDLRRCFYSGDVNA 109
Db 67 DAFGQOQLHLKLODGLFAPGFTLTQV-----GRSPGSEAQHLDPDGLAHCFYSQTVNG 121
QY 110 EPDSFAVSLCGRLGAFYRGAEYVISLPNAS-----APAAQRNCSOGA-----HLQRFG 161
Db 122 DPGSAAALSLCEGVRGAFYLOGEEFFIQAPGAVATRELAPAVPEEBSARPQHILRRR- 180
QY 162 VPGSPGSDPTSRGCVASGNWPAIRALDPYKPRRAGFEGESRRR----- 206
Db 181 ---RRSGGAKCGVMD-----DETLPDTSRSPESQNTNRQWVPVRDPTFODAGKP 226
QY 207 ---SGRAKRFYSIPRYVETLVVADESVMKFKGADLEHYLLTLATAARLYRHPSTILNFI 262
Db 227 SGPGSIRKKREFVSSPRYVETMLVADQSMADFHGSLKHLKLLTLFSAARFYKHPSTRNSI 286
QY 263 NIWVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKLNKVSDDKHPYWDTAIFTTQDLC 322
Db 287 SLVWVKILVYIEQKQPEVTSNAALTLRNFCAWQKLNKVSDDKHPYWDTAIFTTQDLC 346
QY 323 GAFTCDTLGMADVTGMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCVEVFG 382
Db 347 GSHTCDTLGMADVTGMDPKRSCSVIEDDGLQAAFTTAHELGHVFNMPHDNAKHCASLAG 406
QY 383 KLRANHMSPFLIODIRANPWSACSAATITDPLDSHGDCCLLDQPSKPISTLPEDLPAS 442
Db 407 VTGDSHLMASMLSLDHSQSPSCSAIMYTSFLDNGHGECCLMDKPNQIPQDPLPGTSD 466
QY 443 TLSQCELAFGVGSKPCP-YMQYCTKLMCTGKAKGOWQCTRHFPWADGTSCGEGCLK 501
Db 467 DANRQCFTEGEEKKPCDAASTCTLWCTGSGGLLVCTQKHFPWADGTSCGEGKWCVS 526
QY 502 GACVERHNLNKH---RVDGSNAKWDYPGCSRTCGGGVQLARRQCTNTPANGKAYCEGV 558
Db 527 GKCVNKTDM-KHFATPVHGSWGPWGDSCSRTCGGVQVYTMRECDNPVKNKGKCYCEG 585
QY 559 RVKYRSCNLEPCPSASCSFREOCEAFNGYNHSTNRLTLAVAVPKYSGVSPKDKCKL 618
Db 586 RVYRSCNLEDCPDN-NGKTFREOCEAHNEFSKASFGNEPTVEWTPKTAGVSPKDRCKL 644
QY 619 ICRANGTYFVVLAPKVVDGTLCSPDSTSVCGQKCIKAGCDGNLGSKKRFDKCGVCGGD 678
Db 645 TCEAKGIGYFVVLQPKVVDGTPCSPDSTSVCGQCYKAGCDRIIDSKKFDKCGVCGGN 704
QY 679 NKSCKKVTGLTKPMHGNFVVAIPAGASSIDIRQGYKGLIGDDNYALNKSQKYLIN 738
Db 705 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHQRNGSRNNGSFLAIRADGTYILN 764
QY 739 GHFVSAVERDLVKSLLRYSGTGTAVESLOASRPILPLETLVEVLSVCKMTPPVRYYS 798
Db 765 GNFTLSTLEQDLYTKGTVLYRSGSAALEIRSFSPKPELTITQVLMVGHALRPKIKFTY 824
QY 799 YLPKPREDKSHHPKDPGRGPRSVLNSVLSLNOVEQPDPRPPARWAGSWGSPCSASCSG 858
Db 825 FMKKKTES-----FNAIPTFS-----EWIEEWGECCKTCGSG 857
QY 859 LQKRAYDRCRGSAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPCSKSGRGQR 915
Db 858 WQRRVVQCRDINGH---PASECAKEVKPASTRPCADLPCHPMQVGDWSPSCSKTCGKYK 914
QY 916 RSLKCVCHGGRILLARDONLHRKPOE-LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSHESCDPLKKPKHYIDFCTLTQC 950

```
US-10-381-793-3
; Sequence 3, Application US/10381793
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN ADAM-TS-LIKE PROTEIN
; FILE REFERENCE: L10152 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/381,793
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: US 60/235,881
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/XXX,XXX
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-381-793-3

Query Match      48.1%; Score 2485.5; DB 6; Length 951;
Best Local Similarity 48.0%; Pred. No. 2.7e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

QY 1 MLLGILTLAFAGRTAGG--SEPERVWVPIRLDPDINGRRYWRGP-EDSGDQGLIFQI 57
Db 20 LLLASITMLLCARGAHRPTDEDELVP-SLE-----RAPGHDSTTTRL--RL 66
QY 58 TAFQEDFYLHTPDQAFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNA 109
Db 67 DAFGQQLHLKLPDQSGFLAPGFTLTQV-----GRSPGSEAQHLDPDGLAHCFYSGTVNG 121
QY 110 EPDSFAAVSLCGGLRGAFGYGAEVISPLPNAS-----APAAQRNSDGA-----HLLORRG 161
Db 122 DPGSAALSLCEGVRGAYLOGEEFFIQAPGVATERLAPVPEESSAFVQFHILRRR- 180
QY 162 VPGGSGDPTSRGCVASGWNPAIRLALDPYKPRAGFGESRRR----- 206
Db 181 -----RRSGGAKCGVMD-----DETLPDSDRPSQNTNRNQPVRKPTPDAGKP 226
QY 207 -----SGRAKRFVSPRYVETLVVADESMVKFHGADLEHYLLTLLATARLYRHPISLNP 262
Db 227 SGPGSIRKRFVSPRYVETLVVADESMVKFHGADLEHYLLTLLATARLYRHPISLNP 286
QY 263 NIVVVKVLLLRDRDSCPKVTGNAALTRNFCWAKOKLKNVSKDHPYWDTAILFTRODLC 322
Db 287 SLVVKVILVIEEQKGPVTSNAALTRNFCWAKOKLKNVSKDHPYWDTAILFTRODLC 346
QY 323 GATTCDTLGMADVGTWCDPKRSCSVIEDDGLPFAFTAHGLGHVFNPNHNPVNVKCEVFG 382
Db 347 GSHTCDTLGMADVGTWCDPKRSCSVIEDDGLPFAFTAHGLGHVFNPNHNPVNVKCEVFG 406
QY 383 KLRANHMMSPTLTIOTDRANPWSACSAAITDPLDSGHGDCLLDQPSKPISIPEDLPAS 442
Db 407 VTGDSHLMASMLSLDHSQWSPCSAYWVTSFLDNGHGECCMDKPPQIPKLPDLPGLTY 466
QY 443 TISOQELAFAGSKPCP-YMOYCTKLWCTGAKQGVQVOTRHPFWADGTSCEGSKLCKL 501
Db 467 DANRCQOFTFEESKHCPCDAASTCTTLWCVTGSGLLVCQTKHFPWADGTSCEGSKWCVS 526
QY 502 GACVERHNLNKH---RVDGSAKWDPYGPCSRTCCGGVQLARQCTNPTFANGKCYCEGV 558
Db 527 GKCVKNTDM-KHFTATPVHSGMPGPGWDCSRTCCGGVQVYTMRECDNPVKNKGKCYEGK 585
QY 559 RVKRYSCNLEPCPSASGKSPREEQCAFNGYNHSTNRLTLAVAWVPKYSVSPRDKCKL 618
Db 586 RVRYSCNIEDCPDN-NKFTFREQCEAHNEFSKASGNETVETWTPKYAGVSPDKCKL 644
QY 619 ICRANGTYFVYLPKVVVDGTLCSPDSTSVQCGKCIKAGCDNGLGSKKRFDRKCGVCGGD 678
Db 645 TCEARGIGYFVLPKVVVDGTPCSPDSTSVQCGQCVKAGCDRIIDSKKAFDRKCGVCGGN 704
QY 679 NKSCKKYTGTLTKPMHGYNFVAIPAGASSIDIRQRYKGLIGDNDNYLALKNSSQGYLLN 738
```

```
Db 705 GSTCKRMSSGIVTSTRPGYHDIIVIPAGATNIEVKHRNQRNGRNGSFLAIRADGTYLN 764
QY 739 GHFVVSVERDLVVKGSLRLRYSGTGTAVESIQASRPILPTLTVVLSVGRKMTPPRVYSF 798
Db 765 GNFTLSTLEQDLTYKGVLYSGSSAALERIRSFSPKLEPLTTQVLMVGHALRPKIKFTY 824
QY 799 YLPKEPREDKSSHPKDPGRPSVLHNSVLSISNOQVEOPDDRRPPARWVAGSMPGCSASGSG 858
Db 825 FMKKKTES-----FNAITPFS-----EWVTEENGECSTKCGSG 857
QY 859 LOKRAVDCRGSAGORTVPCADAAH--RPVETOAGGE--PCPTWELSAWSPCKSKCGRCFCR 915
Db 858 WQRRVVOCRDINGH---PASECAKEKVPASTRPCADLPCHPQVQVGDWSPCKSKCGRYKK 914
QY 916 RSLKCVGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 915 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 950
```

```
RESULT 13
US-10-369-779-19
; Sequence 19, Application US/10369779
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; APPLICANT: Sharma, Rahul
; APPLICANT: Ramanathan, Chandra
; APPLICANT: Westphal, Ryan
; APPLICANT: Feder, John
; TITLE OF INVENTION: NOVEL HUMAN ADAM-TS POLYPEPTIDE CONTAINING A DISINTEGRIN AND
; TITLE OF INVENTION: METALLOPROTEASE DOMAIN WITH THROMBOSPONDIN TYPE-1 MOTIFS,
; FILE REFERENCE: D0204 NP
; CURRENT APPLICATION NUMBER: US/10/369,779
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/358,151
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-369-779-19
```

Query Match 48.1%; Score 2485.5; DB 6; Length 968;
Best Local Similarity 48.0%; Pred. No. 2.8e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;

```
QY 1 MLLGILTLAFAGRTAGG--SEPERVWVPIRLDPDINGRRYWRGP-EDSGDQGLIFQI 57
Db 37 LLLASITMLLCARGAHRPTDEDELVP-SLE-----RAPGHDSTTTRL--RL 83
QY 58 TAFQEDFYLHTPDQAFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNA 109
Db 84 DAFGQQLHLKLPDQSGFLAPGFTLTQV-----GRSPGSEAQHLDPDGLAHCFYSGTVNG 138
QY 110 EPDSFAAVSLCGGLRGAFGYGAEVISPLPNAS-----APAAQRNSDGA-----HLLORRG 161
Db 139 DPGSAALSLCEGVRGAYLOGEEFFIQAPGVATERLAPVPEESSAFVQFHILRRR- 197
QY 162 VPGGSGDPTSRGCVASGWNPAIRLALDPYKPRAGFGESRRR----- 206
Db 198 -----RRSGGAKCGVMD-----DETLPDSDRPSQNTNRNQPVRDPTPDAGXP 243
QY 207 -----SGRAKRFVSPRYVETLVVADESMVKFHGADLEHYLLTLLATARLYRHPISLNP 262
Db 244 SGPGSIRKRFVSPRYVETLVVADESMVKFHGADLEHYLLTLLATARLYRHPISLNP 303
QY 263 NIVVVKVLLLRDRDSCPKVTGNAALTRNFCWAKOKLKNVSKDHPYWDTAILFTRODLC 322
Db 304 SLVVKVILVIEEQKGPVTSNAALTRNFCWAKOKLKNVSKDHPYWDTAILFTRODLC 363
```

```
QY 323 GATTCDTLGMADVGTMCDDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKVCEVEFG 382
Db 364 GSHTCDTLGMADVGTVCDDSRSCSVIEDDGLQAAFTTAHELGHVFNPHDNNVKAHCSASLNG 423
QY 383 KLRANHMSPTLQIDRANPWSACSAAITDLDSDHSGDCLLDQSPKPSILPELDPGASY 442
Db 424 VTGDSHLMASMLSSLDHSPQSPCSAYWVTSFLDNGHGECLMDKPNPIKLPSPDLPGTLY 483
QY 443 TLSQOCELAFGVGSKPCP-YMQYCTKLWCTGKAGQWVQCTRHPPWADGTSCEGKCLCK 501
Db 484 DANRQCQFTTGEESKHCPCDAASTCTTLWCTGSGLLVCQTKHFPWADGTSCEGKWCVS 543
QY 502 GACVERHNLNKH---RVDGSNAKWDPYGPCSRTCGGVQGLARRQCTNPTFANGKYCEGV 558
Db 544 GKCVNKTDM-KHFATPVHGSWGPWGDSCRTCGGVQVYTMKCDNPPVKNGKYCEGK 602
QY 559 RVKYRSCNLEPCPSSASGKSFREEOCEAFNGYNHSTNRLTLAVANVPKYSVSPDRCKL 618
Db 603 RVKYRSCNIEDCPDN-NGKTFREEOCEAHNEFSKASFGNEPTVTPKYAGVSPKDRCKL 661
QY 619 ICRANGTGYFYVLAPKVVDPGTLCSPDSTSVQVQKICAKGDCNGLGSKKRFKDCGVCGD 678
Db 662 TCEAKGIGYFFVLPQKVVDPGTLCSPDSTSVQVQKICAKGDCNGLGSKKRFKDCGVCGN 721
QY 679 NKSKCKVTGLTKPMHGNFYVVALPAGASSIDIRQYKGLIGDNNYALAKNSOGKYLIN 738
Db 722 GSTCKKMSGIVTSTRPGYHDIVTPAGATNIEVKHNRNGSRNNGSFLAIRAADGTIYN 781
QY 739 GHEVSAVERDLVVKSLRLYSTGTAVESLQASRPLEPLVEVLSVCKMTPPRVYSF 798
Db 782 GNTLTSLQDLTKYKTVLRYSGSSAALRIKSFSPKLEPLTIQVLWGHALRPKIKETY 841
QY 799 YLPEKREDKSHPKDPRGSPVLHNSVLSNQVEQDPRPARVAGSWGKPCSCSGSG 858
Db 842 FMKKTES-----FNAITFS-----ENVIEWGECSKTCGSG 874
QY 859 LQKRAVDCRGSAGQRTVPACDAAH--RPVETQACGP-PCPTWELSAWSPCSKSGRGQR 915
Db 875 WQRRVQOCRDINGH---PASECAKEYKPASTRPCADLPCHQWGVGDWSP-SKTCGKYKK 931
QY 916 RSLKCVGHGRLARQCNLHRKPOB-LDFCVLRPC 950
Db 932 RTLKCVSHDGGVLSNESCDLKKPKHYIDFCILTQC 967

RESULT 14
US-10-391-364-82
; Sequence 82, Application US/10391364
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16558, 14223, 16002, 50566, 50552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MPI03-0190NMN
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
```

```
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining prior Application data removed - See File Wrapper of PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-391-364-82
```

```
Query Match 48.1%; Score 2485.5; DB 6; Length 968;
Best Local Similarity 48.0%; Pred. No. 2.8e-174;
Matches 478; Conservative 159; Mismatches 248; Indels 111; Gaps 22;
```

```
QY 1 MLLLGILTLAFAGTAGG--SEPEREVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFCI 57
Db 37 LLLASITMLLCARGAUGRPTDEELVLP-SLE-----RAPCHDSTTRL--RL 83
QY 58 TAFEDEFYHLHTDPAQFAPAFSTEHLGVPLQGLTGS-----SDLRRCFYSGDVNA 109
Db 84 DAFGQQLHLKLPDPSGLFAPGFTLQTV-----GRSPSEAQLHDPDGLAHCFYSGTNG 138
...QY 110 EPDFAAVSLCGGLRGAFAFYGAIVYISPLPNAS----APAAQNSOGA---HLQRRG 161
Db 139 DPGSAALSLCEGYRGAIFYLOGEEFFTOPAPGVATERLAPAVEEESARPQHILRRR- 197
QY 162 VPGPGSDPTSRCSGVASGNPAILRALDPYKPRRAGGESRRR----- 206
Db 198 -----RRSGGAKCGVMD-----DETLPSTDSRPSQNTNRNQPVRDPTPDAGKP 243
QY 207 ----SGRAKRFVSTPRVETLVVADESVMVPHGADLHLYLLTLATAARLYRHPISILNFI 262
Db 244 SGPGSIRKKRFVSPRYVETMLVADQSMADPHGSLGKHLTLTLFSAARFYKHPISIRSI 303
QY 263 NIVVVKVLLLRDRDSGPKVTGNAALTIRNCAWOKLNKYSDKHPEYWDATAILFTRODLC 322
Db 304 SLVVVKILVIYEEQKGPEVTSNAALTIRNCFNWKQNSPSDRDPEHYDITAILFTRODLC 363
QY 323 GATTCDTLGMADVGTMCDDPKRSCSVIEDDGLPSAFTTAHELGHVFNPHDNNVKVCEVEFG 382
Db 364 GSHTCDTLGMADVGTVCDDSRSCSVIEDDGLQAAFTTAHELGHVFNPHDNNVKAHCSASLNG 423
QY 383 KLRANHMSPTLQIDRANPWSACSAAITDLDSDHSGDCLLDQSPKPSILPELDPGASY 442
Db 424 VTGDSHLMASMLSSLDHSPQSPCSAYWVTSFLDNGHGECLMDKPNPIKLPSPDLPGTLY 483
QY 443 TLSQOCELAFGVGSKPCP-YMQYCTKLWCTGKAGQWVQCTRHPPWADGTSCEGKCLCK 501
Db 484 DANRQCQFTTGEESKHCPCDAASTCTTLWCTGSGLLVCQTKHFPWADGTSCEGKWCVS 543
QY 502 GACVERHNLNKH---RVDGSNAKWDPYGPCSRTCGGVQGLARRQCTNPTFANGKYCEGV 558
Db 544 GKCVNKTDM-KHFATPVHGSWGPWGDSCRTCGGVQVYTMKCDNPPVKNGKYCEGK 602
QY 559 RVKYRSCNLEPCPSSASGKSFREEOCEAFNGYNHSTNRLTLAVANVPKYSVSPDRCKL 618
Db 603 RVKYRSCNIEDCPDN-NGKTFREEOCEAHNEFSKASFGNEPTVTPKYAGVSPKDRCKL 661
QY 619 ICRANGTGYFYVLAPKVVDPGTLCSPDSTSVQVQKICAKGDCNGLGSKKRFKDCGVCGD 678
Db 662 TCEAKGIGYFFVLPQKVVDPGTLCSPDSTSVQVQKICAKGDCNGLGSKKRFKDCGVCGN 721
QY 679 NKSKCKVTGLTKPMHGNFYVVALPAGASSIDIRQYKGLIGDNNYALAKNSOGKYLIN 738
Db 722 GSTCKKMSGIVTSTRPGYHDIVTPAGATNIEVKHNRNGSRNNGSFLAIRAADGTIYN 781
```


QY 739 GHFVSAVERDLVWKSLLRYSGTAVESLOASRPILPTVEVLSVWGKMTFPRVRSF 798
Db 782 GNFTLSTLEODLTYSKTVLRVSGSAALEIRSFSPKEPLTIQVLMVGHALRPKIKFTY 841
QY 799 YLPKEPREKSSHPKDPGRGSPVLHNSVLSNOVEQDDRPAPARVAGSNWGPCSASGSG 858
Db 842 FMKKKTES-----FNAIPTFS-----BWWLHNGECSKTGSG 874
QY 859 LQKRAVDCRGSGAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPSCSKSGRQFGR 915
Db 875 WQRRVVQCRDINGH---PASECAKEVPASTRCPADLPFCPHWQGDWSPSCSKTCGKYKK 931
QY 916 RSLKCVGHGRLRLADQCNLHRKPQE-LDFCVLRPC 950
Db 932 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 967

RESULT 15
US-60-440-068-134
; Sequence 134, Application US/60440068
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CARMAN, JULIE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE
; TITLE OF INVENTION: NF-KB PATHWAY
; FILE REFERENCE: 3053-4191
; CURRENT APPLICATION NUMBER: US/60/440,068
; NUMBER OF SEQ ID NOS: 746
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 134
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-440-068-134

Query Match 48.1%; Score 2482.5; DB 7; Length 967;
Best Local Similarity 48.9%; Pred. No. 4.6e-174;
Matches 485; Conservative 155; Mismatches 250; Indels 101; Gaps 24;

QY 1 MLLGILTLAFAGTAGGSEPEREVVPIRLDPDINGRRYRWGPEDSGOGLFIQTAF 60
Db 36 LLLAALLAVSDALGRPSEDELVVP-ELE-----RAP---GHGTRRLRHAF 81
QY 61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGSSDLRRCFYSGDVNAEPDSF 114
Db 82 DQOLDLELRPDSFLAPGFTLQNGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 136
QY 115 AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQNSQGA---HLLQ---RRGVPGG 165
Db 137 AALSICEGVGAFYLLGEAVFIQPLPAASERLATAAPGKPPAPLQPHLLRRNRQGDVG 196
QY 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPKPRRAGFSGSRSSRS 207
Db 197 TCGVVDEPRPTGKAETDEDEGEDEGPQWS-----PQDPALQGVQP-TGTGS 247
QY 208 GRAKRFVSIPIRYVTLLVVADESVMKFGADLEHYLLTLATAARLYHHPSLNPIINVV 267
Db 248 IRKRRFVSSHRYVTMLVADQSMAEFGSLGKLYLLTFLFSAARLYKHPSIRNSVLVV 307
QY 268 KVLRLDRDSGPKVTGNAALTLRNFCAMOKLNVSKDHPYWDTAILFTRODLGATTC 327
Db 308 KILVIHQKGPVTSNAALTLRNFCNWKQHNPPSDRDAEHYDTALLFIRQDLGCSQTC 367
QY 328 DTLGMADVGTWCDPKRSCSVIEDGLPSAFTTAHELGHVFNMPDHNVKVEEVFGKLRAN 387
Db 368 DTLGMADVGTWCDPSRSCSVIEDGLQAFTTAHELGHVFNMPDHDAKQACASLNGVNDQS 427
QY 388 HMSPTLIQIDRANPNWASCAAITTFDLSHGDCLLDQPSKPLSPEDLPGASYTLISQ 447
Db 428 HMAASMLSNLDHSPWSPSCSAYMITSFIDNGHGECLMDKPHNPQLPGLPGTSDANRQ 487

QY 448 CELAFGVGSKPCP-YMOYCTKLNCTGKAKQOMVQCOTRHPFWADGTSCGBSKLCLKACACVE 506
Db 488 CQTFGEDSKHCPDAASTCTLWCTGTSGVLVCQTKHFPWADGTSCGEGKMGKNGKCVN 547
QY 507 RHLNLKH--RVDSWAKMDPYGPCSRTCGGVQOLARRROCTNPTPANGKCYCEGVVKYK 563
Db 548 KTD-RKHFDTPFHGNWGMGPGWGDSCRTCGGVQOYTWRECDNVPKNGKCYCEGKVKYK 606
QY 564 SCNLEPCPSASGSKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPDRCKLILCRAN 623
Db 607 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLILCOAK 665
QY 624 GTGYEVVYAPKVVDGTLCLSPDSTSVQVQCKIKAGCDGNLGSKKRFDKCGVCGGDNKSK 683
Db 666 GIGYFFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKKFDKCGVCGGNGSTCK 725
QY 684 KVTGLTKPMHGYNFVVAIPAGASSIDIRQYKGLIGDNDYLALKNSSQKYLNLNGHFVV 743
Db 726 KISGSVTSAPKGYHDIITPTGATNIEVKQNRQGRNRNNGSFLAIKAADGTYLNGUYTL 785
QY 744 SAVEDLVVKGSLRLRYSGTAVESLOASRPILPTVEVLSVGMKMTFPRVRSFYLPKE 803
Db 786 STLEQDIMYKVVLYRYSGSAALEIRSFSPKEPLTIQVLTGVGNALRPKIRYTFVKKK 845
QY 804 PREDKSHPKDPRGSPVLHNSVLSNOVEQDDRPAPARVAGSMGSPCSASGSGJQKRA 863
Db 846 -----KES-----FNAIPTFS-----AWVTEEGECSKSELGWQRRL 878
QY 864 VDCRGAGORTVPACDAAH--RPVETOACGE-PCPTWELSAWSPSCSKSGRGRFQKSLRC 920
Db 879 VECRDINGQ---PASECAKEVPASTRCPADHPCPQWQMLGENSSCSKTCGKGYKKSLAC 935
QY 921 VGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db 936 LSHDGGVLSHESCDPLKKPKHYIDFCTMAEC 966

Search completed: May 9, 2003, 15:29:12
Job time : 102 secs

